

GenCore version 5.1.6
copyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 7, 2004, 17:30:20 ; Search time 4482.16 Seconds
(without alignments)
11294.703 Million cell updates/sec

09-245-198A-1

8

gtgctgagcctggcctgg.....ataaatcatgatttctcttc 1168

NTFY NUC

op 10_0 , Gapext 1.0

0272 seqs, 21671516995 residues

s satisfying chosen parameters: 6940544

th: 0

th: 2000000000

nimum Match 0%

ximum Match 100%

string first 45 summaries

nEmbl.*

gb_ba.*

gb_hg.*

gb_in.*

gb_om.*

gb_ov.*

gb_pat.*

gb_ph.*

gb_pl.*

gb_pr.*

gb_ro.*

gb_sy.*

gb_sts.*

gb_un.*

gb_vl.*

em_ba.*

em_fun.*

em_hum.*

em_in.*

em_mu.*

em_om.*

em_ov.*

em_pat.*

em_ph.*

em_pl.*

em_ro.*

em_sts.*

em_un.*

em_vi.*

em_hg_hum.*

em_hg_inv.*

em_hg_other.*

em_hg_mus.*

em_hg_pln.*

em_hg_rtd.*

em_hg_nam.*

em_hg_vrt.*

em_sy.*

em_hgo_hum.*

em_hgo_mus.*

em_hgo_other.*

the number of results predicted by chance to have a

score greater than or equal to the score of the result being prii
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1168	100.0	1168	6	BD062757	BD062757 A t
2	1168	100.0	1239	10	AF030100	AF030100 M
C 3	711	60.9	203083	2	AC069459	AC069459 Mus
C 4	711	60.9	234182	10	AL603707	AL603707 MC
5	628.6	53.8	1353	6	AX201324	AX201324 Sec
6	628.6	53.8	1353	9	AY358870	AY358870 Hon
7	628.6	53.8	1368	9	AF055872	AF055872 Hon
8	628.6	53.8	1421	6	BD090952	BD090952 Apc
9	624	53.4	1306	9	AF030099	AF030099 Hon
10	614.6	52.6	1373	6	BD062758	BD062758 A t
11	597.8	51.2	1236	6	ARI40407	ARI40407 Sec
12	597.8	51.2	1236	6	BD057124	BD057124 Men
13	566.6	48.5	130254	2	AC136195	AC136195 Rat
C 14	566.6	48.5	165316	2	AC119115	AC119115 Rat
15	566.6	48.5	223877	2	AC098923	AC098923 Rat
C 16	566.6	48.5	225077	2	AC136563	AC136563 Rat
17	498.8	42.7	898	6	AX180714	AX180714 Sec
C 18	428	36.6	234801	2	AC118309	AC118309 Rat
19	409.4	35.1	1642	9	BC019047	BC019047 Hon
20	328.4	28.1	1816	9	AY081051	AY081051 Hon
21	304	26.0	218485	2	AC127470	AC127470 Par
C 22	303.8	26.0	60268	9	AC016876	AC016876 Hon
23	278.2	23.8	149555	2	AC126921	AC126921 Bos
24	261	22.3	149736	2	AC126239	AC126239 Fel
25	237.4	20.3	180222	2	AC130192	AC130192 Sus
26	212	18.2	176258	2	AC126925	AC126925 Car
27	111.6	9.6	212093	2	AC126237	AC126237 Car
28	88.4	7.6	7218	6	I66494	I66494 Seque
29	69.8	6.0	195	6	AX379024	AX379024 Seq
30	56.2	4.8	125020	9	AF429315	AF429315 Hon
C 31	55.8	4.8	228056	2	AC144061	AC144061 Mac
C 32	55.8	4.8	303091	2	AC084799	AC084799 Mus
33	55.2	4.7	115145	10	AL807240	AL807240 MC
C 34	55	4.7	185822	2	AC073554	AC073554 Hon
35	54.8	4.7	100511	2	AC010774	AC010774 Hon
36	54.4	4.7	87120	2	AC012225	AC012225 Hon
37	54.2	4.6	191415	10	AC096625	AC096625 Ml
C 38	54.2	4.6	298166	2	AC087563	AC087563 Hon
C 39	53.6	4.6	205350	2	AC078946	AC078946 Mus
C 40	53.2	4.6	49430	2	AC100434	AC100434 Mus
C 41	53.2	4.6	189897	9	AC084064	AC084064 Hon
42	53	4.5	55061	2	AC091597	AC091597 Mus
43	53	4.5	161575	2	AC141041	AC141041 Rat
C 44	53	4.5	187252	10	AL607109	AL607109 Mc
45	52.8	4.5	84514	10	AL627264	AL627264 Mc

ALIGNMENTS

RESULT 1	BD062757	1168 bp	DNA	linear	PAT 27-
LOCUS	BD062757	A tumor necrosis factor related ligand.			
DEFINITION	BD062757				
ACCESSION	BD062757.1	GI:22608360			
VERSION	JP 2001505407-A/1.				
KEYWORDS	unidentified				
SOURCE	unclassified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 1168)				
AUTHORS	Chicheportiche, Y. and Browning, J. L.				
TITLE	A tumor necrosis factor related ligand				
JOURNAL	Patent: JP 2001505407-A 1 24-APR-2001;				
	BIOMED INC, THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GEN				

NF family related

24-APR-2001

07-AUG-1996 US 60/02354

00/040820
VES CHTCHEPOPTCHE TEEBEV I PROBTING

..I2NI5/28,C0/K14/525,G0IN33/68,C0/K16/28
I2NI5/28

A61K39/395, A61K38/19, C07K14/705, C12N15/12

Topology: Linear;

2.6 DS

1. 1168

```
/mol type="genomic DNA"
```

00-1767-32944

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

arity 100.0%; Pred. No. 1.6e-298;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

ATGCTGGGCACCGCTGTCTGCCCCAGGAGCCCTCTCAGGAGGAGCTGACAGCAGAGGA 120

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1. The first step in the process is to identify the problem or issue that needs to be addressed. This involves gathering information and understanding the context of the problem.

2. Once the problem is identified, the next step is to define the objectives and goals of the project. This helps to clarify what needs to be achieved and provides a clear direction for the team.

3. The third step is to develop a plan or strategy to address the problem. This involves breaking down the problem into smaller, manageable tasks and determining the resources needed to complete each task.

4. The fourth step is to implement the plan. This involves putting the strategy into action and monitoring progress regularly to ensure that the project is on track.

5. The final step is to evaluate the results of the project. This involves comparing the actual outcomes against the objectives and goals to determine the effectiveness of the project.

[illegible]

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

GCATTCAGCCCAIATAGAGGTTCATCCTCGGCCAGGACAGGATGGAGCAACAGC 300

AGCTATTGCAGCCCATTAATGAGGGTTCATCCCTCGGCCAGGACAGGATGGAGCAACAAGC 300

'GTGGATGGGACAGTGAGTGGCTGGGAGAGACCAAATCAACAGCTCCAGCCCTCT 360

'GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT 360

TACGACCGCCAGATTGGGGAATTACAGTCATCAGGGCTGGGCTCTACCTGTA 420

TACGACCGCCAGATTGGGGAATTTACAGTCATCAGGGCTGGCTCTAATTACCTGA 420

'CAGGTGCACTTTTGAATGAGCGAAGGGTTCTCTCCCTACCTCCACCTTCCTCCCTCAC

100

[illegible]

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

[illegible]

ATCCGACCCCTCCCTGGGCTCATCTTAAGGCTGGCCCTTCTTAACCTACTTTGG 660

ATCCGCAACCCCTCCCCCTGGGCTCATCTTAAGGCTGCCCCCTTCCTAACCTACTTTGG 660

TTTCAAGTTCACTGAGGGCCTTGCTCTCCAGATTCCTTAAACTTCCCTGGCTC 720

TTTCAAGTTCACTGAGGGCCTTGCTCTCCAGATTCCTTAAACTTCCCTGGCTC 720

100.0%;	Score 1168;	DB 10;	Length 1239;
ity 100.0%;	Pred. No. 1.6e-298;		
nservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 3
AC069459/c

ACCESSION AC069459
 VERSION AC069459.23 GI:14547768
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
 SEQUENCE, 7 unordered pieces.

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chorda
Mammalia: Eutheria: Rodent-

REFERENCE	AUTHORS
I (bases 1 to 203083)	Metzker M L, Lewis J R, Hume J, Edwards C, Harris C

TITLE Direct Submission

REFERENCE 2 (bases 1 to 203083)

TITLE	Direct	Subm
-------	--------	------

JOURNAL

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.ngsc.bcm.edu>
Contact: hase-hel@bcm.tmc.edu

Center project name: MAFO

Sequencing vector: M13; L08821

Chemistry: Dye-terminator Big Dye: 52% of reads

Assembly program: FILIP; version 0.550325
Consensus quality: 272648 bases at least 040

[illegible]

TCCTTGCTCTCTCTTAACATCCCATCCCAACCAACTATCCACCTCACTAGCTCCC 42544

```

1: .23102="Mus musculus"
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_lib="RPCI-23"
misc feature
complement (84050..84131)

```

misc feature

/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."

	60.9%	Score	711	DB	10	Length	234182	
	Identity	99.6%	Pred.	No.	5.7e-177			
	Conservative	0	Mismatches	0	Indels	3	Gaps	3
	GCACATTTTCATCAGGGAAAGCGTGTACCTACCTGAAGCTGGACTTCTCGTGGTGAACGGT	484						
	GCACATTTGATGAGGGAAGCGTGCTACCTTGAGCTGGACTTCTCGTGGTGAACGGT	62664						
	GGCCCTCGCGCTCGCTGCGTGAAGAAATTTCTCAGCCACAGCAGCAAGCTCTCTCTGGGCC	544						
	GGCCCTCGCGCTCGCTGCGTGAAGAAATTTCTCAGCCACAGCAGCAAGCTCTCTCTGGGCC	66204						
	CCGTTGTGTCAGAGTGCTGGGCTGTTCGCGCTGCGGCCAGGGCTCTTCCTTCGG	604						
	CCGTTGTGTCAGAGTGCTGGGCTGTTCGCGCTGCGGCCAGGGCTCTTCCTTCGG	66144						
	CACCCCTCCCTGGGCTCATCTTAAAGGCTGCCCTTCTTAACTATTTTGGACTC	664						
	CACCCCTCCCTGGGCTCATCTTAAAGGCTGCCCTTCTTAACTATTTTGGACTC	66084						
	AGTTCACTGAGGGCGCTTGTCTCTCCAGATTCCTTAACTTTCCCTGGCTCCAGG	724						
	AGTTCACTGAGGGCGCTTGTCTCTCCAGATTCCTTAACTTTCCCTGGCTCCAGG	66024						
	CACACACACTCCCTACCCACCCCACTCTCTCCACCCCTCGCTGCTCTCTGGTC	784						
	CACACACACTCCCTACCCACCCCACTCTCTCCACCCCTCGCTGCTCTCTGGTC	65964						
	TTGT-CTCTCTCAAAGGCAGCAGAGCTTGTTCACATG-TTTTCCATTCCACAGA	842						
	TTGTCTCTCTCTCAAAGGCAGCAGAGCTTGTTCACATGTTTTCATTCCACAGA	65904						
	CCCTTGCTCTTC-TTAACTATCCATCCACACAACTATCACTCACTAGCTCCC	901						
	CCCTTGCTCTTCTTTAACTATCCATCCACACAACTATCACTCACTAGCTCCC	65844						
	CCCTTACTTATCTGACTCCGCCACCCACTCACCCGACACAGTGTTTATTGACTT	961						
	CCCTTACTTATCTGACTCCGCCACCCACTCACCCGACACAGTGTTTATTGACTT	65784						
	ACCAGGCACCTGAGATGGGCTGGACTGTGGTGCAGGAAGCCAGAGAACCCTGGGACT	1021						
	ACCAGGCACCTGAGATGGGCTGGACTGTGGTGCAGGAAGCCAGAGAACCCTGGGACT	65724						
	AGAAGTTCCCACTGTGAGGGGGAAGAGCTGGGGAAGAAGCTCTCTCTGGATCCC	1081						
	AGAAGTTCCCACTGTGAGGGGGAAGAGCTGGGGAAGAAGCTCTCTCTGGATCCC	65864						
	ATTTTTGAAGAAGATCACTATTTTATTATTATTGTGACAAAATGTTAAATGGATTT	1141						
	ATTTTTGAAGAAGATCACTATTTTATTATTATTGTGACAAAATGTTAAATGGATTT	65604						
	GAATAAATCATGATTCCTTC	1168						
	GAATAAATCATGATTCCTTC	65577						

4 1353 bp DNA linear PAT 30-AUG-2001
4 e 3 from Patent WO0153486.
4 4
4.1 GI:15391154
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Homnidae; Homo.

[illegible]

RESULT 9
AF030099
LOCUS
DEFINITION

	722	AGGAGATCACCACACACTCCCTTACCCCACCCCACTCTCTCCACCCCTTC - GGTGG
Dc	898	-----CCTCGACAGCTCTCTGGGCACCCGGTGCCCTCTGCCCCACCCCTCAGCGGT
Qy	781	GSTCAGTCCTGTCTCTCC--TCAAAGGCAGCCAGAGCTTTGTTACATGTTTTCCAA
Dc	953	GCTCCAGACTGCCCTTCCCTCTCTAGAGGCTGCTGGGCTGTTCAAGTGTTTTCCA
Qy	838	-----ACAGACGATTCCTTGCTCTTTTAACAATCCCATCCACACACAACATATCCA
Dc	1013	ACATAAATAFACGATATTCCCACTCTTATCTTACAACCTCCCCCACCGCCCACTCTCCA
Qy	892	ACTAGTCTCCAAAAGCCCTTAC-----TTATCCCTGACTCTCCCCCAACC
Dc	1073	ACTAGTCTCCCAATCCCTGACCCCTTTGAGGCCCCCACTGATCTCGACTCCCCCTCG
Qy	937	CACCGACCAACGCTTTTATTGAATTGTCAC-----
Dc	1133	CAGACCCCGCAGGGCATTTGTTCACCTGTACTCTGTGGCAAGGATGGGTCCAGAAG
Qy	969	-----CAGGCACCTCAGATGGCTGGACCTGGTGGCAGGAAGCAGACGACTGGGA
Dc	1193	CAC TTCAGGCACTAAGAGGGGCTGGACCTGCGCGCAGGAAGCCAAAGACATGGGC
Qy	1024	GCCAGAGTTCCTCCAACTGTGAGGGGGGAAGAGCTGGGGCAAGCTCTCTCCCTTGA--
Dc	1253	GCCAGGAGTTCCTCCAAATGTGAGGGGGCAGA-AACRAGACAAGCTCTCTCCCTTGAGA
Qy	1080	CCTGTGGATTTTGA--AGATACTATTTTATTATTATTGTGACAAAAATGTAA
Dc	1312	CCTGTGGATTTTAAAAACAGATATTTTATTATTATTGTGACAAAAATGTGA
RESULT 11		
LOCUS	AR140407	1236 bp DNA linear PAT 16
DEFINITION	Sequence 1 from patent US 6207642.	
ACCESSION	AR140407	
VERSION	AR140407.1 GI:14482903	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unclassified.	
REFERENCE	1 (bases 1 to 1236)	
AUTHORS	Wiley,S.R.	
TITLE	Member of the TNF family useful for treatment and diagnosis	
JOURNALS	Patent: US 6207642-A 1 27-MAR-2001;	
FEATURES	Location/Qualifiers	
source	1..1236	
	/organism="unknown"	
	/mol_type="unassigned DNA"	
ORIGIN		
Query Match	51.2%; Score 597.8; DB 6; Length 1236;	
Best Local Similarity	75.2%; Pred No. 3e-147;	
Matches 879; Conservative	0; Mismatches 222; Indels 58; G	
Qy	2	GTGCTGACGCTGGGCCCTGGCGCTGGCTGCTGCTGGCTCTCTGCTGGTGTGTGCTGAC
Dc	73	GGCTGTGGGCTGGGCCCTGGCGCTGGCTGCTGCTGGCTCTCTGCTGGCGGTGTGTGCTGAC
Qy	62	GGGAGCTGGGCAACGCTGTCTGCGCAGGAGCCCTTCTCAGGAGAGCTCAGACGACGA
Dc	133	GGGAGCGGGGCATCGCTGTCTGCGCCCAAGGAGCCCTGCCACAGGAGAGTGTGTGTGCGCAG
Qy	122	CGCGGGAGCGCCCTGTAACATGAATCCCAACAGACAGAGAGAAAGCAGGATGTGTGATCC
Dc	193	GAC CAGGACCGCTCGGAACTGAATCCCAACAGAGAGAAAGCAGGATCTGTGGCTGAC
Qy	182	TTGNAACAATAGTTCGGGCTCGAAGAGATGTCTTAAAGGCGGGAAGCGCGCGGCT
Dc	253	CTGAACCGACATAGTTCGGGCTCGAAGAGATGTACCTTAAAGGCGGGAACACAGCGGC

301 TATTGCAGCCCATTTATGAGGTTTCATCTCTGGCCGACAGGACAGGATGGAGCACAAAGCA
 372 GATCGCAGCCCATTTATGAAGTTTCATCCACGACCTGGACAGGACGCGCAGGCA
 361 GGATGGACAGTGTGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCTG
 432 GGACGGACAGTGTGCTGGGAGAGCCAGAAATCAACAGCTCCAGCCCTCTG
 421 CGACCCGAGATTGGGGAAATTACAGTCAATCAGGGCTGGGCTCTACTACCTGTAC
 492 CAACCGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTCTACTACCTGTAC
 481 GGTGCATTTGATGAGGGAAGGCTGTCTACTGAAGCTGACTTTGCTGTGTGAAC
 552 GGTGCATTTGATGAGGGAAGGCTGTCTACTGAAGCTGACTTTGCTGTGTGAAT
 541 GCTGGCCTGCTGCTGGAGAAATCTCAGCCACACAGCAGCAAGCTCTCTGGG
 612 GCTGGCCTGCTGCTGGAGAAATCTCAGCCACTGCGGCGAGTTCCCTCGGG
 601 AGCTCCGTTGTGTCAGAGTGTCTGGGCTGTTCGCGCTGCGGCCAGGGTCTTGCCCT
 672 GCTCGGCTCTGCCAGAGTGTCTGGGCTGTTCGCGCTGCGGCCAGGGTCTCTGCTG
 661 CCGCACCTCTCCCTGGGCTCATCTTAAGGCTGCCCCCTTCTTAACCTACTTTGGA
 732 CCGCACCTCTCCCTGGGCTCATCTCAAGGCTGCCCCCTTCTCTCACTACTTTGGA
 721 TCAAGTTTCACTGAGGGGCTGTGCTCTCCAGATCTCTTAAATTTTCCTGCTGCTCC
 792 CCAGGTTCACTGAGGGGCTGTGCTCTCCGCGAGTGTGTCAGGGGTGCGGCTCC
 780 CATCACACACCTCTCCCTACCCACCCACCTCTCCACCCGCTC-GTGTGCTCTT
 847 CCTCGACAGCTCTCTGGGCAACCGGTCCCTCTGCCCCACCTCAGCCGCTCTTT
 837 AGTCTGTCTCC--TCAAAGGACGACAGCTGTGTTCATGTTTCCATTC--
 907 AGACTGCCCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTGTTTTCATTC--
 891 -ACAGAGTATCTCTGCTCTTTAAATCCATCCACACCAATATCCACCTC
 967 AATACAGTATTCCTCTTATCTTACAAACACCCACCGCCCACTCTCCACCTC
 936 CTCCCAAGCCCTAC-----TTATCCTGACTCCCCACCCCACT
 1027 CTCCCCAATCCTGACCTTTGAGGCCCCGAGTATCTGACTCCCCCTCGGCA
 968 GACCAGCTGTTTATGACTTTGTGCAC-----
 1087 CCCCAGGGCATGTGTTCACTGTACTCTGTGGCAAGGATGGTCCAGAGACCC
 1023 CAGCACTGAGATGGGCTGGACCTGTGGGAGGAGCCAGAGAACCTGGGACTAG
 1147 CAGCACTAAGAGGGGCTGACCTGTGGGCGCAGGAAGCCAAAGAGACTGGGCTAG
 1079 AAGTTCCTCAACTGTGAGGGGAAGAGCTGGGGAAGAGCTCTCCCTGGA----TC
 1206 GAGTTCCTCAATGTGAGGGGCGAGA-AACAAGACAAGCTCTCTCCCTTGAGAAATC
 1108 GGATTTTGAAGAATCTATTTTT
 1235 GGATTTTAAACAGATATTTT

34 1236 bp DNA linear PAT 27-AUG-2002
 of the tnf family useful for treatment and diagnosis of
 3.
 34 4.1 GI:22602730

[illegible]

06:25:09 2004

us-09-245-198a-1.rge

TCGACACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCTCACCTACTCTCGA 732
TTCAAGTTCACTGAGGGGCTTGTCTCTCCAGATTCTTAACTTTCCCTGGCTCC 721
TCCAGTTCACTGAGGGGCTTGTCTCTCCAGATTCTTAACTTTCCCTGGCTCC 792
GCATCACACACCTCCCTTACCCACCCCTCTCTCCACCTCTCTCCACCTCTCTCT 780
-CCTCGACAGCTCTCTGGGACCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 847
CAGTCT 837
CAGACCTGCT 907
--ACAGACCT 891
AAATACAGTATCT 967
GCTCCCAAGGCT 936
GCTCCCAAGGCT 1027
CGACCT 968
CCCCCAGGCT 1087
-CAGGCT 1023
TCAGGCT 1147
GAGTCT 1079
GAGTCT 1206
TGGATTTTAAAGATCTATTTT 1108
TGGATTTTAAAGATCTATTTT 1235

95 130254 bp DNA linear HTG 06-JUN-2003
norvegicus clone RP31-258K6 strain Brown Norway, WORKING
SEQUENCE, 12 ordered pieces.

95.3 GI:31442440
TGTS_PHASE2; HTGS_DRAFT.
norvegicus (Norway rat)
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
uses 1 to 130254)
llis A., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
ley R.W., Bouffard G.G., Brinkley C., Brooks S., Cariaga K.,
Coleman B., Coleman H., Engle J., Granite S., Guan X.,
J., Haghighi P., Han J., Hansen N., Ho S.-L., Hu P.,
J., Idol J.R., Karlins E., Kwong P., Laric P., Lee-Lin S.-Q.,
i.R., Maduro Q.L., Maduro V.B., Margulies E.H., Masiello C.,
i.B., McDowell J., Paguirigan C., Pearson R., Portnoy M.E.,
A., Reddix-Dugue N., Schandler K., Schueler M.G., Shah K.,
C., Stantripop S., Thomas J.W., Thomas P.J., Tsipouri V.,
i.L., Wetherby K.D., Wiggins L., Young A. and Green E.D.
Comparative Sequencing Initiative
uses 1 to 130254)
E.D.
Submission
ted (30-OCT-2002) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
uses 1 to 130254)
E.D.
Submission

Submission
ted (30-OCT-2002) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
uses 1 to 130254)
E.D.
Submission

Submitted (06-JUN-2003) NIH Intramural Sequencing Center,
Grovmont Circle, Gaithersburg, MD 20877, USA
On Jun 6, 2003 this sequence version replaced gi:27753660.
Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: dcf
Center clone name: 258K06

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicat
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associat
with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 128169 bases at least Q40
Consensus quality: 128674 bases at least Q30
Insert size: 150000; agarose-fp
Insert size: 129154; sum-of-contigs
Quality coverage: 10.78x in Q20 bases; agarose-fp
Quality coverage: 12.52x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that ha
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 10521: contig of 10521 bp in length
* 10522 10621: gap of unknown length
* 10622 13327: contig of 2706 bp in length
* 13328 13427: gap of unknown length
* 13428 28924: contig of 15497 bp in length
* 28925 29024: gap of unknown length
* 29025 39201: contig of 10177 bp in length
* 39202 39301: gap of unknown length
* 39302 41906: contig of 2605 bp in length
* 41907 42006: gap of unknown length
* 42007 70195: contig of 28089 bp in length
* 70196 77661: contig of 7366 bp in length
* 77662 94161: gap of unknown length
* 94162 94261: gap of unknown length
* 94262 97982: contig of 3721 bp in length
* 97983 98082: gap of unknown length
* 98083 105400: contig of 7318 bp in length
* 105401 105500: gap of unknown length
* 105501 127855: contig of 22355 bp in length
* 127856 127956: gap of unknown length
* 127956 130254: contig of 2299 bp in length.

Location/Qualifiers
1. 130254
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Brown Norway"

FEATURES
source

AAAGATAAATCATGATTCTTCTTC 111224

3 223877 bp DNA linear HTG 10-MAY-2003
norvegicus clone CH230-154B15, WORKING DRAFT SEQUENCE, 3
ed pieces.

3 3.8 GI:30521223

CS_PHASE1: HTGS_DRAFT: HTGS_FULLTOP.

norvegicus (Norway rat)

na; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

ses 1 to 223877)

Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
Howells, S., Huly, S., Hume, J., Idlebird, D., Jackson, A.,
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Loulseg, H., Lozano, R.J., Lu, X., Ma, J.,
Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mapua, P., Martin, K., Martin, R., Martinez, E.,
McLeod, M.P., McNeill, T.Z., Meenen, E.,
Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Neal, D., Newton, N., Nguyen, N., Norris, S.,
Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rockey, T., Rojao, A., Rose, M., Rose, R., Ruiz, S.J.,
Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D.,
Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Zhang, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
hausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Gibbs, R.A.

Submission

ished

ses 1 to 223877)

K.C.

Submission

ted (06-NOV-2001) Human Genome Sequencing Center, Department
ecular and Human Genetics, Baylor College of Medicine, One

REFERENCE AUTHORS TITLE JOURNAL

COMMENT

Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 223877)
Rat Genome Sequencing Consortium.

Submitted (10-MAY-2003) Human Genome Sequencing Center, De
of Molecular and Human Genetics, Baylor College of Medicine
Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:25008075
The sequence in this assembly is a combination of BAC bases
and whole genome shotgun sequencing reads assembled using
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig
in the feature table below represents a scaffold in the A
assembly (a 'contig-scaffold'). Within each contig-scaffold
individual sequence contigs are ordered and oriented, and
by sized gaps filled with Ns to the estimated size. The se
may extend beyond the ends of the clone and there may be s
contigs within a contig-scaffold that consist entirely of
genome shotgun sequence reads. Both end sequences and whol
shotgun sequence only contigs will be indicated in the fea
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GIOK

Center clone name: CH230-154B15

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 214785 bases at least Q40

Consensus quality: 216908 bases at least Q30

Consensus quality: 218593 bases at least Q20

Estimated insert size: 227169; sum-of-contigs estimati

Quality coverage: 7x in Q20 bases; sum-of-contigs esti

* NOTE: Estimated insert size may differ from sequence len

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_da

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 221327: contig of 221327 bp in length

* 221328 221427: gap of unknown length

* 221428 222652: contig of 1225 bp in length

* 222653 222752: gap of unknown length

* 222753 223877: contig of 1125 bp in length.

FEATURES

Location/Qualifiers

1..223877

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-154B15"

1..1822

/note="wgs contig"

complement(217607..218056)

/note="clone boundary"

clone_end:T7

site:EcoRI

end_sequence:RWBB008TJB"

misc_feature

misc_feature

complement(217607..218056)

note="clone boundary"

clone_end:T7

site:EcoRI

end_sequence:RWBB008TJB"

ORIGIN

Query Match 48.5%; Score 566.6; DB 2; Length 223877;

Best Local Similarity 89.1%; Pred. No. 1.1e-138;

Matches 669; Conservative 0; Mismatches 74; Indels 8; G

Qy 425 CAGTGCACCTTTCATGAGGAAAGGCTGTCTACTGAAGCTGGACTTGTGCTGAA

Db 147649 CAGTGCACCTTTCATGAGGAAAGGCTGTCTACTGAAGCTGGACTTGTGCTGAA

06:25:09 2004

us-09-245-198a-1.rge

CTGGCCCTGCGCTGCTGGAAGAAATCTCAGCCACAGCAGCAAGCTCTCTCTGGGCCC 544
|||||
CTGGCCCTGCGCTGCTGGAAGAAATCTCAGCCACAGCAGCAGCTCTCTGGGCCC 147768
|||||
CTCGCTTTGCGCCAGGCTGCTGGGCTGTTGCCGTGGCGCCAGGCTCTCCCTTCGG 604
|||||
CTCGCTGTTGCGCCAGGCTGCTGGGCTGTTGCCGTGGCGCCAGGCTCTCCCTTCGG 147828
|||||
CGCACCTCCCTCGGCTCATCTTAAAGCTGCCCCCTTCTTAACCTACTTTGACTC 664
|||||
CGTACCTCCCTCGGCTCATCTTAAAGCTGCCCCCTTCTTACCTACTTTGACTC 147888
|||||
CAAGTTCACTGAGGGGCTGCTCTCCAGATTCTTTAAACTTTCC ---TGGCTC 720
|||||
CAAGTTCACTGAGGGGCTGCTCTCCAGATTCTTTAAACTTTAACTTAACAGCTC 147948
|||||
CAGCATCACACACCTCCCTACCCACCCCTCTCCACCCCTCGCTGCTCCTT 780
|||||
CAGCACACAGACCTCCACCCACCCAGCCCTCTCCACCTCTAGCTGTTCTT 148008
|||||
CAGTCCCTGCTCTCTTCCAGGCGAGCCAGGCTTGTGACATGTTTCCATTCCA 838
|||||
CAGTCCCTGCTCTCTTCCAGGCGAGCCAGGCTTGTGACATGTTTCCATTCCA 148068
|||||
ACGTATCCTGCTCTTCT-TAACATCCCATCCACACACATATCCACCTCACTAGC 897
|||||
ACATATTCCTGCTCTTCTGTAGCTCCACCCACCCAGCTGCGCGCTCAGAGC 148128
|||||
CAAGCCCTTACTTATCCCTGACTCCCCACCCACTCACCCGACACGTTTATTG 957
|||||
CAAGCCCTTACTTATCCCTGACTCCCCACCCACTCACCCGACACGTTTATTG 148187
|||||
TTGTGACCCAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAACTGG 1017
|||||
ATGTGATCAGGCACTCAAGATAGGCTGGACCTGGTGGCAGGAAGCCAGAACTGG 148247
|||||
CAGGCCAGAGTTCCCACTGTGAGGGGAGAGCTGGGGACAAGCTCTCCCTGGA 1077
|||||
TAGGCCAGAGTTCCCAATGTGAGGGGAAAGAGATGGAGACAAGCTCTCCCTGGA 148307
|||||
TGTGGATTTGAAAAGATACTATTTTATTATTATTGTGACAAAATGTTAAATGGA 1137
|||||
TGTGGATTTGAAAAGATACTATTTTATTATTATTGTGACAAAATGTTAAATGGA 148367
|||||
TAAAGAGATAAATCATGATTTCTTTC 1168
|||||
TAAAGAGATAAATCATGATTTCTTTC 148398
|||||

April 8, 2004, 02:01:42
; secs

6:25:11 2004

us-09-245-198a-1.rst

GenCore version 5.1.6
copyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 7, 2004, 17:30:19 ; Search time 2831.52 Seconds
(without alignments)
12318.149 Million cell updates/sec

09-245-198A-1

8
gtgctgagcctggcctgg.....ataaatcatgattctcttc 1168

NTITY NUC

op 10.0 , Gapext 1.0

13289 seqs, 14931090276 residues

s satisfying chosen parameters: 55026578

th: 0

th: 2000000000

nimum Match 0%

string first 45 summaries

T:*

em_estba:*

em_esthum:*

em_estin:*

em_estmu:*

em_estov:*

em_estpl:*

em_estro:*

em_htc:*

gb_est1:*

gb_est2:*

gb_htc:*

gb_est3:*

gb_est4:*

gb_est5:*

em_estfun:*

em_estom:*

em_gss_hum:*

em_gss_inv:*

em_gss_pln:*

em_gss_vrt:*

em_gss_fun:*

em_gss_nam:*

em_gss_mus:*

em_gss_pro:*

em_gss_rod:*

em_gss_phg:*

em_gss_vrl:*

gb_gss1:*

gb_gss2:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry	ch	Length	DB	ID	Description
.7	1033	11	AK020909	AK020909	Mus muscu
.0	892	14	CB204861	CB204861	AGENCOURT
.6	918	10	BF577781	BF577781	602092080
.8	665	13	BY742288	BY742288	BY742288

5	549	47.0	2237	11	AK044387	AK044387 MU
6	533.4	45.7	543	29	CG565104	CG565104 OS
7	519.4	44.5	731	12	BI871711	BI871711 60
8	510	43.7	554	29	CG629394	CG629394 OS
9	507.4	43.4	728	12	BI870393	BI870393 60
10	504	43.2	561	10	AW763237	AW763237 UR
11	497	42.6	533	10	BE628951	BE628951 UU
12	489.6	41.9	650	12	BG404836	BG404836 60
13	488.8	41.8	687	13	BQ208433	BQ208433 UI
14	481.2	41.2	584	10	AW917574	AW917574 ES
15	480.8	41.2	939	14	CB849011	CB849011 MR
16	480.4	41.1	777	12	BI819200	BI819200 60
17	476	40.8	482	29	CG653257	CG653257 OS
18	473	40.5	834	12	BI766766	BI766766 60
19	467.8	40.1	948	13	BQ707185	BQ707185 AG
20	464.8	39.8	666	14	CF126539	CF126539 UI
21	456.4	39.1	828	12	BI596681	BI596681 60
22	448.2	38.4	621	29	CG584545	CG584545 OS
23	440.6	37.7	545	14	CB141389	CB141389 X-
24	437.8	37.5	471	9	AA221610	AA221610 my1
25	430.2	36.8	498	29	CG554711	CG554711 OS
26	426.8	36.5	624	13	BU759448	BU759448 UI
27	422.6	36.2	697	14	CF126932	CF126932 UI
28	421.6	36.1	474	29	CG609156	CG609156 OS
29	421.6	36.1	963	13	BQ671259	BQ671259 AG
30	420	36.0	483	29	CG646715	CG646715 OS
31	419.2	35.9	940	13	BQ884231	BQ884231 AC
32	410.4	35.1	456	13	EX634398	EX634398 BX
33	406.8	34.8	471	29	CG568080	CG568080 OE
34	406.4	34.8	1071	12	BM921213	BM921213 AC
35	400.4	34.3	418	29	CG611020	CG611020 OE
36	393.2	33.7	494	29	CG596702	CG596702 OS
37	390.6	33.4	445	9	AA870722	AA870722 VQ2
38	390.2	33.4	493	29	CG498076	CG498076 OS
39	386.6	33.1	531	29	CG590009	CG590009 OS
40	383.2	32.8	468	29	CG573612	CG573612 OS
41	366.6	31.4	951	13	BQ674188	BQ674188 AG
42	360	30.8	360	10	BE654876	BE654876 UI
43	353.2	30.2	483	29	CG525153	CG525153 OS
44	353	30.2	405	9	AJ854476	AJ854476 UI-
45	350.6	30.0	414	29	CG599845	CG599845 OS

ALIGNMENTS

RESULT 1	AK020909	1033 bp	mRNA	linear	HTC 20-
LOCUS	Mus musculus adult retina cDNA, RIKEN full-length enriched				
DEFINITION	clone:A930030D13 product:tumor necrosis factor (ligand)				
	superfamily, member 12., full insert sequence.				
ACCESSION	AK020909.1	GI:12861640			
VERSION	AK020909.1	GI:12861640			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murina				
REFERENCE	1				
AUTHORS	Carninci,P. and Hayashizaki,Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K				
	Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs				
JOURNAL	prepare full-length cDNA libraries for rapid discovery of n				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
	11042159				

Procurement: Jeffrey E. Green, M.D.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: InCyte Genomics, Inc.
Distribution: InCyte Genomics, Inc.
Through the I.M.A.G.E. Consortium/LLNL at:
/image.llnl.gov

LLM9767 row: p column: 20
quality sequence start: 17
quality sequence stop: 724.
Location/Qualifiers
1. 918
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4206595"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP Co24"
/note="Organ: Colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

53.6%; Score 625.8; DB 10; Length 918;
arity 99.2%; Pred. No. 4.7e-152;
conservative 0; Mismatches 2; Indels 3; Gaps 3

GCCTGCTTGGCCCTCCTGCTGGTGTGGTTCAGGCTGGGAGCTGGGCAACGCTGTC 81
SCCTGCTTGGCCCTCCTGCTGGTGTGGTTCAGGCTGGGAGCTGGGCAACGCTGTC 60
CAGAGGCTTCTCAGGAGGAGCTGACAGCAGAGGACCGCGGAGCCCTCGAACT 141
CAGGAGCTTCTCAGGAGGAGCTGACAGCAGAGGACCGCGGAGCCCTCGAACT 120
CCCCAGACAGAGAAAGCCAGGATGTGTACCTTCTTGGACAACTAGTCCGGCC 201
CCCCAGACAGAGAAAGCCAGGATGTGTACCTTCTTGGACAACTAGTCCGGCC 180
AGAAGTGCTCCTAAAGCCCGGAAGCGCGGCTTCGCGAGCTATTGCAGCCCAATTA 261
AGAAGTGCTCCTAAAGCCCGGAAGCGCGG-CTCGCGAGCTATTGCAGCCCAATTA 239
GTTTCATCTTCGGCCAGCAGAGATGGAGCACACAGCAGGTGTGATGGGACAGTGAG 321
GTTTCATCTTCGGCCAGGACAGGATGGAGCACACAGCAGGTGTGATGGGACAGTGAG 299
TGGGAGAGACACAAATCAACAGCTCCAGCCCTCTCGCTTACGCGCCAGATGG 381
TGGGAGAGACACAAATCAACAGCTCCAGCCCTCTCGCTTACGCGCCAGATGG 359
TTTACAGTCATCAGGGCTGGGCTCTACTACTGTACTGTACGTGCACCTTTCATGA 441
TTTACAGTCATCAGGGCTGGGCTCTACTACTGTACTGTACGTGCACCTTTCATGA 419
AAGGCTGTCTACTGAAGCTGGACTTGTGTTGGTGAACGGTGTGCTGGCCCTTCGCTG 501
AAGGCTGTCTACTGAAGCTGGACTTGTGTTGGTGAACGGTGTGCTGGCCCTTCGCTG 479
GAAGAAATCTCAGCCACAGCAGCAAGCTCTCCTGGGCCCCAGCTCCGTTTGTGCCA 561
GAAGAAATCTCAGCCACAGCAGCAAGCTCTCCTGGGCCCCAGCTCCGTTTGTGCCA 539
TCTGGGCTGTTCGGCTGGCGGCCA-GGTCTTCCTTCGGATCGCACCTCCCGCTG 621
TCTGGGCTGTTCGGCTGGCGGCCA-GGTCTTCCTTCGGATCGCACCTCCCGCTG 598
CATCTTAAGGCTGCGCCCTTCCTAAACCTACTTTGGACCTCTTCAAGTTCACCTGAGG 681
CATCTTAAAGGCTG-CGCTTCTCTAACTAACTTTGGACTCTTCAAGTTCACCTGAGG 657

Location/Qualifiers

134

1

2

3

11

10

—

2

RESULT 5

LOCUS

.....

ACCESSION

KEYWORDS

ORGANIS

AUTHORS

JOURNAL

PUBMEI

AUTHORS

TITLE

JOURNAL

MEDLINE
PUBMED

REFERENCES

TITLE

JOURNAL
OF
MEDICAL
INQUIRY

REFERENCES

AUTHORS

TITLE

REFERENCE

NTCOM Consortium and the RIKEN Genome Exploration Research
 Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 770 full-length cDNAs
 420, 563-573 (2002)
 (see 1 to 2237)
 J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Iida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 H., Kawahara, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Kurihara, K., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nara, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohse, N.,
 Ii, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 S., Saeki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
 tsu, M. and Hayashizaki, Y.
 Submission

admission (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Medical and Chemical Research (RIKEN), Laboratory for Genome Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Yokohama 226-0045, Japan (E-mail: genome-res@riken.go.jp, <http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222, -45-503-9216)

Library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken C Sciences Center and Genome Science Laboratory in RIKEN. One of Experimental Animal Research in Riken contributed to the mouse tissues.

RNA was provided by Dr. Stefano Gustinich (Department of Physiology, Harvard Medical School, 220 Longwood Ave., Boston, MA USA) whose assistance is gratefully acknowledged. Please see our web site for further details.

<http://genome.gsc.riken.go.jp/>
<http://fantom.gsc.riken.go.jp/>

```

cp: /mnt/ncg1/3801/ncg50.gp/
Location/Qualifiers
1. .2237
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:A930010H24"
/db_xref="MGI:2459696"
/db_xref="taxon:10090"
/clone="A930010H24"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
346..1578
/notes="unnamed protein product; putative
tumor necrosis factor (ligand) superfamily, member 12.
(MGD MGI:1916833, GB|NM_023517, evidence: BLASTN, 99%,
match=787)"

```

```

match=7677
/codon_start=1
/protein_id="BAC31897.1"
/db_xref="GI:26336424"
/translation="MAARSQRRRGRRGPGTALLAPVLESLGLACLGLLVVVVSL
GSNAQLQAQEPSEOEETAEPRPEPEINLPDSEODVYPLEQLVPRRSAPGKGA
RPRRAIAHYEVHVRPGDQAQAGVDGTVSGWEEYKINSSPLEDRQIGETVIRAG
LYYICQVHPDEGKAVYLLDLLVNGVLALRLEEFSAASSPGPQLRLCQTELQSL
RREYSRLRGSGSGQKQERPMQSLWESPDLEAKWDGAKRRRRLVTKRKKHS
LVLVPPVNTISKASDVTEVMVQPVLRGRGLQAQGDIVRVMDTGIIYLLSQVLFHD
VTFMGVYVRREGQSGRRRTLFRICRSMPPDPORAYNSCYSGAVFHLHQGDIITVKIPRA
NAKLISPHGHFELGFKVL"
2215..2220
/note="putative"
2237
/note="putative"

```

47.0%; Score 549; DB 11; Length 2237;
Parity 99.5%; Pred. No. 8.2e-132;

Matches	562;	Conservative	0;	Mismatches	0; Indels	3;
1	GGTGCTGAGCCTGGGCGCTGGCGCTGCCTTCCTGGCCTCCTGCTGGTCTGGTGTCT	QY				
417	GGTGCTGAGCCTGGGCGCTGGCGCTGGCGCTGCCTTCCTGGCCTCCTGCTGGTCTGGTGTCT	Db				
61	GGGGAGCTGGGCAACGCTGTCTGGCC---CAGGAGGCTTCTCAGGAGAGCTGACAC	QY				
477	GGGGAGCTGGGCAACGCTGTCTGTGCCCAGCAGGAGGCTTCTCAGGAGGAGCTTGACAC	Db				
118	GGACCGCGGGAGCCCCCTGAACTGAATCCCCAGACAGAGAAAGCCAGGAGTGTG	QY				
537	GGACCGCGGGAGCCCCCTGAACTGAATCCCCAGACAGAGAAAGCCAGGAGTGTG	Db				
178	TTTCTTGGAAACAATACTAGTCGGGCTCGAAGAAAGTCTCTCTAAAGCCGGGAAGGGCG	QY				
597	TTTCTTGGAAACAATACTAGTCGGGCTCGAAGAAAGTCTCTCTAAAGCCGGGAAGGGCG	Db				
238	TCGCCAGCTATTGCAGGCCCATATTAGGTTTCATCTCGGCCCGACAGCAGGATGAC	QY				
657	TCGCCAGGAGCTATTGCAGGCCCATATTAGGTTTCATCTCGGCCCGACAGCAGGATGAC	Db				
298	AGCAGGTGTGGATGGGACACAGTCAGTGGCTGGGAAAGAGCAAAAATCAACAGCTCTCT	QY				
717	AGCAGGTGTGGATGGGACACAGTCAGTGGCTGGGAAAGAGCAAAAATCAACAGCTCTCT	Db				
358	CTCGCGCTACGACCGCGCAGATTGGGGAATTTACATCATCAGGGCTGGGGCTTCTACT	QY				
777	CTCGCGCTACGACCGCGCAGATTGGGGAATTTACATCATCAGGGCTGGGGCTTCTACT	Db				
418	GTACTGTCAAGTGACATTTGATCAGGGAAGGCTCTTACTGAGCTGGAGCTGGACTTGG	QY				
837	GTACTGTCAAGTGACATTTGATCAGGGAAGGCTCTTACTGAGCTGGAGCTGGACTTGG	Db				
478	GAAACGGTGTCTGGCCCTCGGCTGCTGGAAAGAAATTCCTCAGCCACAGCAGCAAGCT	QY				
897	GAAACGGTGTCTGGCCCTCGGCTGCTGGAAAGAAATTCCTCAGCCACAGCAGCAAGCT	Db				
538	TGGGCCCCAGCTCCGTTTGTGCCAG	QY				
957	TGGGCCCCAGCTCCGTTTGTGCCAG	b				

RESULT 6					
CGS65104					GSS 0:
LOCUS	CGS65104	543 bp	DNA	linear	
DEFINITION	Ost189654 Mus musculus 129Sv/Ev Mus genomic clone OST189654, genomic survey sequence.				

ACCESSION	CG565104
VERSION	CG565104.1
	GI:37351691

KEYWORDS GSS.

SOURCE
Mus musculus (house mouse)

ORGANISM *Mus musculus*

Eukaryota; Metazoa; Chorda

Mammalia; Eutheria; Rodent

REFERENCE 1 (bases 1 to 543)

AUTHORS
Zambrowicz, B. P., Abuin, A.,

Piggott, J., BeltrandelRio, M.

Friddle, C.J., Gupta, A., Ha

Key, B.W. Jr., Klipp, P., Koh-

Payne, R., Potter, D.G., Qian

Sparks, M.J., Van Slightenho

Zhu, Q.; person, C. and sand
Wnt1 kinase deficiency low

TITLE WHK1 KINASE DEFICIENCY LOWERS SCREEN TO IDENTIFY POTENTIAL

JOURNAL OF THE
ACADEMY OF
SCIENCE AND
TECHNOLOGY
OF AMERICA

JOURNAL
 PLOC: NALL: ACAD: SCI: U.S
 Contact: Zambrowicz BP
 COMMENT

COMMENT: ZAMBROWICZ BF
OmniBank

Lexicon Genetics Incorporated
OMNIBANK

LEXICON GENETICS INCORPORATED
4000 Research Forest Drive

4000 Research Triangle Drive
Email: materials@lexgen.com

Gene trap sequence tag gene

described in Zambrowicz et

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

06:25:11 2004

us-09-245-198a-1.rst

TTAGGGGCCCTGGTCTCCCGAGTGTGCTCCAGGCTGCCGCTCC 708

94 554 bp DNA linear GSS 02-OCT-2003
515 Mus musculus 129Sv/Ev Mus musculus genomic clone
515, genomic survey sequence.

94
94.1 GI:37453243

sculus (house mouse)

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ses 1 to 554)
wicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
t,J., Beltrando,R.H., Buxton,E.C., Edwards,J., Finch,R.A.,
e,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
i,M.J., Van Slichthorst,I., Vogel,P., Walke,W., Xu,N.,
Person,C. and Sands,A.T.
inase deficiency lowers blood pressure in mice: a gene-trap
i to identify potential targets for therapeutic intervention
Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
t: Zambrowicz Bp
unk

in Genetics Incorporated
Research Forest Drive, The Woodlands, TX 77381, USA
materials@lexgen.com

rap sequence tag generated by 3' RACE from mouse ES cells as
bed in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Gene Trap.

Location/Qualifiers

1..554
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129Sv/Ev"
/db_xref="taxon:10090"
/clone="OST341515"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

43.7%; Score 510; DB 29; Length 554;
arity 98.9%; Pred. No. 6.8e-122;
onservative 0; Mismatches 3; Indels 3; Gaps 3;

TTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT 360
GTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCT 65

TTAGCACCCAGATTTGGGAATTTACGTATCATCGGGCTGGGCTTACTACTGTA 420
TTAGCACCCAGATTTGGGAATTTACGTATCATCGGGCTGGGCTTACTACTGTA 125

TCAGGTGCATTTTGATGAGGAAAGGCTGTCTACCTGAAGCTGGACTTGTGGTGA 480
TCAGGTGCATTTTGATGAGGAAAGGCTGTCTACCTGAAGCTGGACTTGTGGTGA 185

TTGTG-CTGGCCCTGCGCTGCTGGAAGAAATTTCTAGCCACAGCAGCAAGCTTCCTG 539
TTGTGCTTGGCCCTGCGCTGCTGGAAGAAATTTCTAGCCACAGCAGCAAGCTTCCTG 245

CCAGCTCCCTTTTGTGACAGGTGTCTGGGCTGTGCGCTGCGCCAGGGCTTCCC 599
CCAGCTCCCTTTTGTGACAGGTGTGCGCTGCGCTGCGCCAGGGCTTCCC 305

AGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCTTCTTAACCTACTTTG 659
AGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCTTCTTAACCTACTTTG 365

QY 660 GACTCTTTCAAGTTCACTAGGGGCTTGTCTCTCCAGATTCTTAAACTTTCCCT
Db 366 GACTCTTTCAAGTTCACTAGGGGCTTGTCTCTCCAGATTCTTAAACTTTCCCT

QY 720 CCAGGAGATCACACACCTCCCTACCCACCCACCCACCTCCACCCCTCGCTG
Db 426 CCAGGAGATCACACACCTCCCTACCCACCCACCCACCTCCACCCCTCGCTG

QY 780 TGGTCCAGTCTCTGT-CTCTCTCAAAGGAGCCAGAGCTTGTTCACATG-TTTCG
Db 486 TGGTCCAGTCTCTGTCTCTCAAAGGAGCCAGAGCTTGTTCACATGTTTTCG

QY 838 ACAGACGTA 846
Db 546 ACAGACGTA 554

RESULT 9

BI870393

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

AP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Gene Index
 ished (1997)
 t: Robert Strausberg, Ph.D.
 cgapbs-r@mail.nih.gov
 lone is available royalty-free through LLNL; contact the
 Consortium (info@image.llnl.gov) for further information.
 83048
 imer: -40RP from Gibco
 uality sequence stop: 437.
 Location/Qualifiers
 1. .533
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:3373444"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"
 /clone_lib="Soares mammary gland NMLMG"
 /note="Vector: pTV730-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pTV73 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Fatima Bernaldo."

42.6%; Score 497; DB 10; Length 533;

arity 99.4%; Pred. No. 1.7e-118;

onservative 0; Mismatches 0; Indels 3; Gaps 3;

GTGCTGCCCTCGCTCGTGGAGAAATTCACGCCAGCAGCAAGCTCTCTGG 540
 |||||
 GTGCTGGCCCTCGCTCGTGGAGAAATTCACGCCAGCAGCAAGCTCTCTGG 60
 |||||
 CAGCTCCGTTTGTGCCAGTGTCTGGGCTGTTCGGCTGCGCCAGGCTCTTCCT 600
 |||||
 CAGCTCCGTTTGTGCCAGTGTCTGGGCTGTTCGGCTGCGCCAGGCTCTTCCT 120
 |||||
 ATCCGACCCCTCCCTCGCTCACTTAAAGCTGCGCCCTTCTTAACTACTTTGG 660
 |||||
 ATCCGACCCCTCCCTCGCTCACTTAAAGCTGCGCCCTTCTTAACTACTTTGG 180
 |||||
 TTTCAGTTTCACTGAGGGGCTTGTCTCCAGATTCCTTAACTTTCCCTGGCTC 720
 |||||
 TTTCAGTTTCACTGAGGGGCTTGTCTCCAGATTCCTTAACTTTCCCTGGCTC 240
 |||||
 AGCATCACACACCTCCCTACCCACCCCACTCTCTCCACCCCTCGCTGCTCTT 780
 |||||
 AGCATCACACACCTCCCTACCCACCCCACTCTCTCCACCCCTCGCTGCTCTT 300
 |||||
 CAGTCT 838
 |||||
 CAGTCT 360
 |||||
 CGTATCCTTGTCT 897
 |||||
 CGTATCCTTGTCT 420
 |||||
 AAAGCCCTTACTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 957
 |||||
 AAAGCCCTTACTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
 |||||
 TGTGACACGAGCACTGAGATGGGCTGGACCTGGTGGAGAGCCAGAG 1010
 |||||
 TGTGACACGAGCACTGAGATGGGCTGGACCTGGTGGAGAGCCAGAG 533
 |||||

BG404836
 LOCUS
 DEFINITION 60242016QF1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4527C
 mRNA sequence.
 BG404836
 BG404836.1 GI:13298284
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information ca
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM10435 row: h column: 15
 High quality sequence stop: 468.
 Location/Qualifiers
 1. .650
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:4527038"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: 1
 Site 2: SalI; Cloned unidirectionally; oligo-dT 1
 Average insert size 3.3 Kb. Library enriched for
 full-length clones and constructed by Life Techn
 Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 41.9%; Score 489.6; DB 12; Length 650;
 Best Local Similarity 92.0%; Pred. No. 1.6e-116;
 Matches 550; Conservative 0; Mismatches 44; Indels 4;
 QY 291 GAGCACAGCAGGTGTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAAC
 DB 25 GACCCCATTCAGGTGTGGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAAC
 QY 351 CCAGCCCTCTCGCTAGCAGCCGAGATTTGGGAAATTTACAGTCATCAGGGCTGG
 DB 85 CCAGCCCTCTCGCTAGCAGCCGAGATTTGGGAAATTTACAGTCATCAGGGCTGG
 QY 411 ACTACCTGTACTGTTCAGTGCACCTTTGATGAGGAAAGGCTGTCTACCTGAAGCTC
 DB 145 ACTACCTGTACTGTTCAGTGCACCTTTGATGAGGAAAGGCTGTCTACCTGAAGCTC
 QY 471 TGCTGGTGAACGGTGTGTCTGGCCCTCGCTGGCTGCCTGGGAAGAAATTTCTCAGCCACAGC
 DB 205 TGCTGGTGAACGGTGTGTCTGGCCCTCGCTGGCTGCCTGGGAAGAAATTTCTCAGCCACAGC
 QY 531 GCTCTCCCTGGGCCCGCAGCTCCGTTTGTGCGCAGGTGTCTGGGCTGTGGCGCTGGCG
 DB 265 GCTCTCCCTGGGCCCGCAGCTCCGTTTGTGCGCAGGTGTCTGGGCTGTGGCGCTGGCG
 QY 591 GGTCTTCCTTCGATCGCAGCCCTCCCTCGGCTCATCTTAAAGCTGCCCCCTTC
 DB 324 GGTCTTCCTTCGATCGCAGCCCTCCCTCGGCTCATCTTAAAGCTGCCCCCTTC
 QY 651 CCTACTTTGGACTCTTTCAAGTTTCACTGAGGGGCTTGTCTCTCCAGATTCTTTA
 DB 384 CCTACTTTGGACTCTTTCAAGTTTCACTGAGGGGCTTGTCTCTCCAGATTCTTTA

GGCTCCAGGAGCATCACACACCTCCCTACCCACCCCACTCTCCACCCCTC 770
 |||||
 GGCTCCAGGAGCATCACACACCTGACTACCCACACAGTAATCCACCCACTC 503
 |||||
 TCCTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
 |||||
 TCCTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563
 |||||
 TCCACAGAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
 |||||
 TCCACAGAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619
 |||||
 TCCACAGAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

3 687 bp mRNA linear EST 02-MAY-2002
 9-coc-1-22-0-UI.s1 UI-R-EPO Rattus norvegicus cDNA clone
 9-coc-1-22-0-UI 3', mRNA sequence.

3.1 GI:20424898

norvegicus (Norway rat)

ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

es 1 to 687)

,M.F., Lennon,G. and Soares,M.B.

zation and subtraction: two approaches to facilitate gene
 ry
 Res. 6 (9), 791-806 (1996)

Soares, MB

ated Laboratory for Computational Genomics

ity of Iowa

ton Road , 4156 MEBRF, Iowa City, IA 52242, USA

9 335 8250

9 335 9565

bento-soares@iowa.edu
 nence contained an oligo-dT track that was present in the
 cleotide that was used to prime the synthesis of first
 cDNA and therefore this may represent a bonafide poly A
 the sequence tag present in the cDNA between the NotI site
 oligo-dT track served to identify it as a clone from the
 red duodenum library cDNA library Preparation: M.B. Soares
 ne distribution: clones will be available through Research
 s (www.resgen.com)
 mer: M13 Forward
 es.

Location/Qualifiers

1. 687

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-EPO-coc-1-22-0-UI"

/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/clone_lib="UI-R-EPO"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; UI-R-EPO is a
 subtracted cDNA library constructed according to Bonaldo,
 Lennon and Soares, Genome Research, 6:791-806, 1996. First
 strand cDNA synthesis was primed with an oligo-dT primer
 containing a Not I site. Double stranded cDNA was ligated
 to an EcoR I adaptor, digested with Not I, and cloned
 directionally into pT73-Pac vector. The oligonucleotide
 used to prime the synthesis of first-strand cDNA contains
 a library tag sequence that is located between the Not I
 site and the (dT)18 tail. The sequence tags for this
 library are: distal colon, GAAAGTGTCTCC; osteoblast,

ORIGIN

Query Match 41.8%; Score 488.8; DB 13; Length 687;
 Best Local Similarity 87.9%; Pred. No. 2.6e-116;
 Matches 590; Conservative 0; Mismatches 73; Indels 8; G;

QY 496 GGGCTGCTGGAGAAATTCAGCCACAGCAGCAGCTCTCTGGGCCCCAGTCC
 |||||
 DB 687 GGGCTGCTGGAGAAATTCAGCCACAGCAGCAGCTCTCTGGGCCCCAGTCC

QY 556 GTGCCAGGTGCTGGGCTGTCCGCTGCGCCAGGGTCTCCCTTCGGATCCGCA
 |||||
 DB 627 GTGCCAGGTGCTGGGCTGTCCGCTGCGCCAGGGTCTCCCTTCGGATCCGTA

QY 616 CCCCTGGGCTCATCTTAAGGCTGCCCTTCCTTAACCTACTTTGGACTCTTTCAAG
 |||||
 DB 567 CCCCTGGGCTCATCTTAAGGCTGCCCTTCCTTAACCTACTTTGGACTCTTTCAAG

QY 676 CTGAGGGGCTTGCTCTCCAGATTCCTTAAACTTTCCCTCCCTCCAGAGC
 |||||
 DB 507 CTGAGGGGCTTGCTCTCCAGATTCCTTAAACTTTCCCTCCCTCCAGAGC

QY 732 CCACACCTCCCTACCCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCAG
 |||||
 DB 447 CAGCAGCTCCCTACCCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCAG

QY 792 GT-CTCTCTCAAGGAGCAGCAGCTTGTTCACATG-TTTCATTCCACAGAGCT
 |||||
 DB 387 GTCTCTCTCAAGGAGCAGCAGCTTGTTCACATG-TTTCATTCCACAGAGCT

QY 850 TTGCTCTTCT-TAACATCCATCCACCACTATCCACCTCAGCTAGCTCCCAAA
 |||||
 DB 327 CCGTTCTTCTGTAACCTCCACCCACAGCTGCCCGCTCCAGAGTCCCAAA

QY 909 CTACTTATCCCTGACTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTTGT
 |||||
 DB 267 CCACGTA-CCCGACTACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTTGT

QY 969 CAGGCACTCAGATGGCTGGACCTGGTGGCAGGAGCCAGAGAACCTGGGACTAGG
 |||||
 DB 208 CAGGCACTCAGATGGCTGGACCTGGTGGCAGGAGCCAGAGAACCTGGGACTAGG

QY 1029 AGTTCCCACTGTGAGGGGAGAGCTGGGGAGCAAGCTCCTCCCTGGATCCCTGT
 |||||
 DB 148 AGTTCCCACTGTGAGGGGAGAGCTGGGGAGCAAGCTCCTCCCTGGATCCCTGT

QY 1089 TTTGAAAGATCTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA
 |||||
 DB 88 TTTGAAAGATCTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAA

QY 1149 ATAAATCATGA 1159
 |||||
 DB 28 ATAAATCATGA 18

RESULT 14
 AW917574
 LOCUS

DEFINITION
 EST348878 Rat gene index, normalized rat, norvegicus, Bent
 Rattus norvegicus cDNA clone RGIEP49 5' end, mRNA sequence

AW917574
 VERSION
 AW917574.1 GI:8083328
 EST.

ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin

ses 1 to 584)
H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
age, A.R. and Adams, M.D.
name Project: Generation of a Rat EST (REBT) Catalog & Rat
index
ished (1998)
t: Lee, NH
stitute for Genomic Research
Medical Center Drive, Rockville, MD 20850, USA
301)-838-3529
301)-838-0208
nhlee@tigr.org
lone is available through the ATCC, contact the ATCC
3-365-2700 for further information
imer: M13 Reverse.
Location/Qualifiers
1..584
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="RGIEF49"
/tissue_type="mix - brain, ovary, placenta, kidney, lung,
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/clone_lib="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; Estimated insert size approx.1 kb"
41.2%; Score 481.2; DB 10; Length 584;
arity 91.3%; Pred. No. 2.3e-114;
onservative 0; Mismatches 46; Indels 5; Gaps 2;
GGCCAGGACAGGATGGAGCAACAGCAGGTGGATGGGACAGTGGCTGGGAA 331
AGCCAGGACAGGATGGAGCAACAGCAGGTTGGATGGGACAGTGGCTGGGAA 61
CCAAATCAACAGCTCCAGCCCTCTGCCTACGACCGGCAGATTGGGAATTTACA 391
CCAAATCAACAGCTCCAGCCCTCTGCCTATGACCGCCAGATTGGGAATTTAG 121
TCAGGGCTGGGCTCTACTACCTGACTGTCTAGGTGCACTTTGTATGAGGAAAGGCT 451
TCAGGGCTGGGCTCTACTACCTGACTGTCTAGGTGCACTTTGTATGAGGAAAGGCA 181
ACCTGAGCTGGACTGTCTGTGTAACGGTGTCTGGCCCTCGCTGCTGCTGGAAGAA 511
ACCTGAGCTGGACTGTCTGTGTAACGGTGTCTGGCCCTCGCTGCTGCTGGAAGAA 241
CAGCCACAGCAGCAAGCTCTCTGGGCCCCAGCTCCGTTTGTGCGAGGTGTCTGGG 571
CAGCCACAGCAGCAGGCTCTCTGGGCCCCAGCTCCGTTTGTGCGAGGTGTCTGGG 301
TGCCGCTGGGCGCAGGCTCTCTGGGCCCCAGCTCCGTTTGTGCGAGGTGTCTGGG 631
TGCCGCTGGGCGCAGGCTCTCTGGGCCCCAGCTCCGTTTGTGCGAGGTGTCTGGG 361
CTGCCCCCTTCTTAACCTTCTGAGCTCTTCAAGTTCAGTTCAGTGGGCGCTTGTCTC 691
CTGCCCCCTTCTTACCTTCTGAGCTCTTCAAGTTCAGTTCAGTGGGCGCTTGTCTC 421
AGATTCTTAACTTTCCCTCTGGGCCCCAGCTCCGTTTGTGCGAGGTGTCTGGG 747
AGATTCTTAACTTTACCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 481
XCCACTCTCTCAACCCCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 807
XCCACTCTCTCAACCCCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 541
XAGAGCTTGTTCATATG-TTTCATTTCCACAGCAGTATCC 849

Db 542 AGCCAGNGCTTGATGACATGTTTTTCCATTCCACAGACATATTC 584
RESULT 15
CB849011
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
1 (bases 1 to 939)
Yu, J., Farjo, R., MacNee, S.P., Baehr, W., Stambolian, D.E. an
Swaroop, A.
Annotation and analysis of 10,000 expressed sequence tags
developing mouse eye and adult retina
Genome Biol. 4 (10), R65 (2003)
2281944
14519200
Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kellogg Eye Center, University of Michigan
540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu.
Location/Qualifiers
1..939
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="retina"
/clone_lib="MOUSE ADULT RETINA"
/note="Vector: pSPORT1"
ORIGIN
Query Match 41.2%; Score 480.8; DB 14; Length 939;
Beat Local Similarity 90.4%; Pred. No. 3.7e-114;
Matches 553; Conservative 0; Mismatches 52; Indels 7;
QY 225 GGAAGGCGCGCCCTCGCGAGCTATTCAGCCCATTTATGAGGTTCATCTCGGCCF
DB 25 GGAAGGCTGTGCTCGCTCAGGTACCGGTCCGGAATTCGCCGCTCGACCCAGCGG
QY 285 AGGATGGAGCACACAGGCTGTGGATGGACAGTGGCTGGGAGAGAGACCAAF
DB 85 AGGATGGAGCACACAGGCTGTGGATGGACAGTGGCTGGGAGAGAGACCAAF
QY 345 ACAGCTCCAGCCCTCTCGCTTACGACCCAGATTGGGGAATTTACAGTCATCAG
DB 145 ACAGCTCCAGCCCTCTCGCTTACGACCCAGATTGGGGAATTTACAGTCATCAG
QY 405 GGCTCTACTACTGTACTGTTCAGGTGCACTTTGATGAGGGAAGGCTCTTACCTC
DB 205 GGCTCTACTACTGTACTGTTCAGGTGCACTTTGATGAGGGAAGGCTCTTACCTC
QY 465 TGGACTTGTCTGTCGACGCTGTGTGCGCTGCGCTGCGTGGAGAAATTTCTCAGCC
DB 265 TGGACTTGTCTGTCGACGCTGTGTGCGCTGCGCTGCGTGGAGAAATTTCTCAGCC
QY 525 CAGCAAGCTCTCTGGGCCCCAGCTCCGTTTGTGCGAGGTGTCTGGGCTGTTCGCC
DB 325 CAGCAAGCTCTCTGGGCCCCAGCTCCGTTTGTGCGAGGTGTCTGGGCTGTTCGCC
QY 585 GGCAGAGGCTCTTCCCTTCGGATTCGACCCCTCCCTCGGCTCATCTTAAAGCTGC
DB 385 GGCAGAGGCTCTTCCCTTCGGATTCGACCCCTCCCTCGGCTCATCTTAAAGCTGC
QY 645 TCCTAACCTACTTTGGATCTCTTTCAAGTTCACTGAGGGGCGCTTGTCTCTCCAGAT

6:25:11 2004

us-09-245-198a-1.rst

P

ACCTACTTTGGACTCTTTCAAGTTCACTGAGGGGCTTGTCTCTCCAGATACCTT 504
TTCCCTGGCTCC-AGGAGCATCACACACCTCCCTAGCCACCCCTCTCTCCA 763
TTNCTGGCTCCAAAGGAGATCACACACCTGCTTACCCACCCCTCTCTGCA 564
T-CGCTGCTCCTTGGTCCAGTCC--TGTCTCTCTCTCAAGGCA--GCCAGAGCT 817
TNNCTGCTCCTTGGGACCANTCCTGTTTCTCTCTCTNCTNAAAGGCANNNCAANAGCT 624
ACATGTT 829
ACATGTT 636

Pril 7, 2004, 23:15:14
secs

16:25:10 2004

us-09-245-198a-1.rng

GenCore version 5.1.6
copyright (c) 1993 - 2004 CompuGen Ltd.

c search, using sw model

il 7, 2004, 17:30:19 ; Search time 446.331 Seconds
(without alignments)
11117.066 Million cell updates/sec

09-245-198A-1
8
gtgctgagccggcgctgg.....ataaatcatgattctcttc 1168

NTITY_NUC

op 10.0 , Gapext 1.0

3863 seqs, 2124099041 residues

s satisfying chosen parameters: 6747726

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
sting first 45 summaries

Geneseq_29Jan04: *
geneseqn1980s: *
geneseqn1990s: *
geneseqn2000s: *
geneseqn2001as: *
geneseqn2001bs: *
geneseqn2002s: *
geneseqn2003as: *
geneseqn2003bs: *
geneseqn2003cs: *
geneseqn2004s: *

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ch	Length	DB	ID	Description
.0	1168	2	AAV18599	Mus muscu
.0	1239	9	ADC97713	Murine FL
.9	701	2	AAx23425	Mouse TNR
.8	1353	3	AAA49717	Human PRO
.8	1353	6	ABK40255	CDNA enco
.8	1421	2	AAx56000	Human tum
.4	1306	7	ACC57587	Polynucle
.4	1306	7	ACC57901	Human TWE
.4	1306	9	ADC35205	Human CDN
.9	1364	6	ABK34881	Human CDN
.6	1373	2	AAV18600	Homo sapi
.2	1236	2	AAV47613	TNF relat
.2	1236	4	AAD04350	Human TRE
.8	1030	2	AAx23424	Human TNR
.7	898	4	AAS03964	Expressio
.5	412	9	ADB56326	Toxicity-
.2	408	7	ABX37032	Bovine ES
.1	282	2	AAT22190	Human gen
.0	195	6	ABK29540	Colon ade
.6	493	8	ACH34013	Human end
.6	65	6	ABN55975	Mouse spl
.2	1064	6	ABT09678	Human PAL
.9	264	7	ABx52254	Bovine ES

C	24	52.6	4.5	2000	7	ADA71938	Ric
C	25	46.8	4.0	1117	9	ADC86688	Hur
C	26	46.6	4.0	3133	9	ADC86738	Hur
	27	46.6	4.0	53522	6	AAD30228	Hur
	28	46.6	4.0	53526	2	AAT94101	Hur
	29	46.6	4.0	53577	2	AAT18551	Hur
	30	46.6	4.0	53577	2	AAT94108	Hur
	31	45.8	3.9	1065	6	ABT09682	Hur
	32	45.8	3.9	105325	6	ABK94407	DNI
C	33	45.6	3.9	5452	9	ADC86736	Hur
	34	45	3.9	1337	2	AAC17263	Hur
C	35	44.8	3.8	1000	3	AAA02484	Hur
C	36	44.8	3.8	3163	9	ADC87060	Hur
C	37	44.8	3.8	12733	6	ABK98631	Vec
C	38	44.8	3.8	12733	8	ACD13882	L.
C	39	44.8	3.8	12739	6	ABK98592	Vec
C	40	44.8	3.8	12739	8	ACD13843	Plz
C	41	43.6	3.7	16167	6	ABL33083	Hur
C	42	43.6	3.7	16167	6	ABL34529	Hur
C	43	43.6	3.7	16167	6	ABL70254	Chc
	44	43.2	3.7	324	6	ABL51826	Sei
	45	43.2	3.7	720	4	AAH05001	Hur

ALIGNMENTS

RESULT 1
AAV18599
ID AAV18599 standard; cDNA; 1168 BP.
AC
XX AAV18599;
DT 21-JUL-1998 (first entry)
XX
DE Mus musculus tumour necrosis factor related ligand (TRELL) gene.
XX
KW TRELL; tumour necrosis factor related ligand; tnf; treatment; can
KW autoimmune disease; immune system; stimulation; suppression;
KW graft rejection; ds.
XX
OS Mus musculus.
FH
FH Key Location/Qualifiers
FT CDS 2..679
FT /tag= a
FT /note= "tumour necrosis factor related ligand"
XX
PN W09805783-A1.
XX
PD 12-FEB-1998.
XX
PF 07-AUG-1997; 97WO-US013945.
XX
PR 07-AUG-1996; 96US-0023541P.
PR 18-OCT-1996; 96US-0028515P.
PR 18-MAR-1997; 97US-0040820P.
XX
PA (BIOJ) BIOGEN INC.
PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
PI Chicpeportiche Y, Browning JL;
XX
DR WPI; 1998-145619/13.
DR P-PSDB; AAW47524.
XX
PT Tumour necrosis factor related ligand - useful for, e.g. treating
PT auto-immune disease and immune responses to tissue grafts.
XX
PS Claim 2; Page 45-46; 69pp; English.
XX
CC The sequence is that encoding mouse tumour necrosis factor relat
(TRELL). TRELL or active fragments can be included with a carrier

n describes isolated Tumor Necrosis Factor (TNF) family peptides: APO4, APO6, APO8 and APO9 or their active fragments. APO4 is useful for diagnosing prostate cancer by levels of APO4 in an individual. Prostate cancer can also be APO4 selective binding agents linked to a therapeutic disease are also useful for identifying selective s, useful in diagnosis/treatment of disease by binding of peptide/active fragment which is extracellular, or the cell surface. The binding is preferably performed in peptides/active fragments are also useful for screening and antagonists by binding and observing the change in APO4 active pharmacological agents useful in diagnosis or disease are also identified using APO4 polypeptides/active APO4 signal transducer molecules that specifically interact asic domain of APO4 and detecting a change in level of APO4 method is performed in vivo or in vitro. APO polypeptides as immunogens for preparing antibodies. APO4 is also agnosis/treatment of developmental or gestational i. APO8 was transfected to human breast carcinoma cell line induced apoptosis

BP; 139 A; 210 C; 203 G; 149 T; 0 U; 0 Other;

arity 59.9%; Score 699.4; DB 2; Length 701;
onservative 0; Mismatches 1; Indels 0; Gaps 0;

TCGTGGTTCAGCTGGGAGCTGGCAAGCTGTCTGCCAGGAGCTTCTCAGGAG 103
TCGTGGTTCAGCTGGGAGCTGGCAAGCTGTCTGCCAGGAGCTTCTCAGGAG 60
TGACAGCAGAGACCGCGGAGCCCTGAACTGATCCCGACAGAGAGAAAGC 163
TGACAGCAGAGACCGCGGAGCCCTGAACTGATCCCGACAGAGAGAAAGC 120
ATGTGGTACCTTTCTTGGAACAACTAGTCGGGCTCGAAGAGTGTCTCTAAAGGC 223
ATGTGGTACCTTTCTTGGAACAACTAGTCGGGCTCGAAGAGTGTCTCTAAAGGC 180
AGGCGCGGCTTCGCGAGCTATTGACAGCCCAATTATGAGGTTCACTCTGGCCAGGA 283
AGGCGCGGCTTCGCGAGCTATTGACAGCCCAATTATGAGGTTCACTCTGGCCAGGA 240
ATGAGCACAAGCAGGTGTGATGGACAGTGCAGTGGCTGGGAGAGACCAAAATC 343
ATGAGCACAAGCAGGTGTGATGGACAGTGCAGTGGCTGGGAGAGACCAAAATC 300
AGTCCAGGCTTCGCGTACGACCGCGAGATTGGGGAATTTACAGTCACTCAGGCT 403
AGTCCAGGCTTCGCGTACGACCGCGAGATTGGGGAATTTACAGTCACTCAGGCT 360
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 463
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 420
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 523
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 480
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 583
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 540
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 643
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 600
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 703
TCTACTACTGTCTGAGTGCAGTGCAGTGTGATGGGAAAGGCTCTCTACCTGAAG 660

QY 704 TAAACTTTCCCTGGCTCCAGGAGCATCACCACCTCCCTA 744
Db 661 TAAACTTTCCCTGGCTCCAGGAGCATCACCACCTCCCTA 701
RESULT 4
AAA49717
ID AAA49717 standard; cDNA; 1353 BP.
XX
AC AAA49717;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human PRO207 cDNA clone DNA30879-1152.
XX
KW PRO207; human; antitumor; tumour; therapy; cytostatic; breast c;
KW ovarian cancer; renal cancer; colorectal cancer; uterine cancer;
KW prostate cancer; lung cancer; bladder cancer;
KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 58..807
FT sig_peptide /*tag= a
FT 58..177 /*tag= b
FT mat_peptide 178..804 /*tag= c
XX
PN WO200037638-A2.
XX
PD 29-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US028565.
XX
PR 22-DEC-1998; 98US-0113296P.
PR 08-MAR-1999; 99WO-US005028.
PR 21-APR-1999; 99US-0130232P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-0134287P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA
PI Napier MA, Pitti RM, Wood WI;
XX
DR WPI; 2000-442668/38.
DR P-PSDB; AAY95338.
XX
PT Novel composition to inhibit neoplastic cell growth or for treat:
PT in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866
XX
PS Claim 20; Fig 3; 172pp; English.
XX
CC The present sequence is that of cDNA clone DNA30879-1152 (ATCC 2;
CC encoding human PRO207 (see AAY95338), which shows homology to se
CC members of the tumour necrosis factor family, especially human
CC lymphotoxin (23.4%). The cDNA was identified in a foetal kidney c
CC library following identification of an expressed sequence tag wi
CC homology to human Apo-2 ligand. A claimed method for inhibiting
CC growth of a tumour cell comprises exposing the tumor cell to PRO
CC PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301, PRO526,
CC PRO356, PRO509 or PRO866 (see AAY95337-49), their agonists or ch
CC polypeptides incorporating them. The tumour is especially a cance
CC selected from breast, ovarian, renal, colorectal, uterine, prost
CC lung, bladder and central nervous system cancer, melanoma and lei
CC Nucleic acids encoding PRO179 etc. are used in the recombinant p:

our polypeptides

BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

53.8%;	Score 628.6;	DB 3;	Length 1353;
ity 76.2%;	Pred. No. 1.2e-159;		
nservative	0;	Mismatches 219;	Indels 73;
			Gaps 10;

GAGCTGGGCGTGCGCTGGCCCTGCTTGGCCCTCTCTGGTCTGGTCTGAGCTG 61
GGGCTTGGGCGCTGGCGCTGGCCCTGCTCGGCTCTCTGGCCGTGGTCTGAGTTTG 189
CTGGGCGAACGGTGTCTGCCAGAGCGCTTCTCAGGAGGAGCTGACAGCAGAGAC 121
CCGGGCATCGTGTTCGCCCCAGGAGCGCTGCCAGGAGGAGCTGGTGGCAGAGAG 249
GGAGCCCGCTTGAACTGAATCTCCACAGACAGAGGAAAGCCAGGATGTGTGTAACCTTTC 181
GGACCCGCTCGGAACTGAATCTCCACAGACAGAGGAAAGCCAGGATCTTCGCGCCCTTTC 309

ACAACTAGTCGGCCTCGAAGAAAGTGTCTCTAAAGCCGGAAGCGCGGCTTCG 241
 CCGACTAGTTTCGGCCTCGAGAAGTGCACCTAAAGCCGGAATACACGGGCTCGA 369
 TATTGCAGCCCATTTATGAGGTTTCATCTCGGCCAGGACAGGATGGAGCAACAAGCA 301
 GATCGCAGCCCATTTATGAGTTTCATCCAGCACTTGAGCAGGAACGAGCGCAGGCA 429

'GGATGGGACAGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCTG 361
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
'GGACGGGACAGTGAGTGGCTGGGAGGAGCAGAGAATCAACAGCTCCAGCCCTCTG 489

CGACCGCCAGATTGGGGAATTACAGTCTATCAGGGCTGGGCTCTACTACCTGTAC 421
GACCGGCGGTCATCGGGGAGCTTTATAGTCTCCCGGCTGGGCTCTACTACCTGTAC 549

[illegible]

GCTGGCCCTGCGCTGCCTGGAAGAAATCTCAGCCACAGCAGCAAGCTCTCCTGGG 541

GCTCCGTTTGTGCCAGGTGTCTGGGCTGTTCGCTGCGCGCCAGGGTCTTCCCTT 601

CCGCACCCCTCCCTGGGCTCATCTTAAGGCTGCCCCCTTCTTAACCTACTTTGGA 661

CCGCAACCCCCTCCCCCGGGGCCCATCTCAGGGCGTGCCTCTCCAGATTCTTAAACTTCCTCCGTGGCTCC 721

CCAGGTTCACTGAGGGGCTTGGTCTCCCGGAGTCGTCCAGGCTGCCGGGCTCC 849
;CATCACCAACCTCCCTACCCCAACCCGCACTCCTCCACCCCTC-GCTGCTCCTT 780

.CCTCGACAGCTCTGTGGGCACCCGGTCCCTCTGCCCCACCCCTCAGCCGCTCTTT 904
 :AGTCCTGTCTCTCC--TCAAAGGCAGCCAGAGCTTGTTCACATGTTTCCATTCC- 837

2AGACCTGCCCCCTCCCTCTAGAGGCTGCCTGGGCCCTGTTACAGTGTTTTCCATCCC 964
 --ACAGAGGTATCCCTTGCTCTTCTTAACATCCCATCCACCAACTATCCACCTC 891

AAATACAGTATCCCGACTCTTATCTTACAACATCCCCCAGCGCCCACTCTCCACCTC 1024
 3CTCCCAAAGCCCTAC-----TTATCCTGACTCCCCCAGCCCACT 936

3CTCCCCAATCCCTGACCTTTGAGGCCCCAGTGATCTGACTCCCCCTGGCCA 1084
2GACCACGTTGTTATTGACTTTGTGCAC----- 968

1085	DB	CAGACCCCAGGTCATGTGTTCATCTGTACTCTGTGGCAGAGATGGGTCCAGAG
969	QY	-----CAGGCACTCAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGAACTCTGGGA
1145	DB	CAC TTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGGC
1024	QY	GCAGAGATTTCCCAACTGTGAGGGGGAAGAGCTGGGNCAGAGCTCCTCCCTGGA
1205	DB	GCCAGGAGTTCCCAAAATGTGAGGGGCGAGA--AACAGACAAGCTCTCCCTTTCAGAG
1080	QY	CCTGTGATTTTCAAA--AGATACTATTTTTATTATTATTGTGCACAAAATGT---TJ
1264	DB	CCTGTGATTTTTAAACAGATATATTTTTATTATTATTGTGCACAAAATGTTGAT
1135	QY	GGATATTAAAGAGAAATAATCATGA 1159
1324	DB	GGATATTAAATAGAATAAGTCATAA 1348

RESULT 5
ABK40255
ID ABK40255 standard: cDNA: 1353 BP.

XX	
AC	ABK40255;
XX	
DT	15-JUL-2002 (first entry)

XX cDNA encoding human PRO207 polypeptide.
DE
XX
XX Human, p90; benign tumour; malignant tumour.

KW leukaemia; neuronal disorder; stromal disorder; blastocoelec diso.
 KW inflammatory disorder; immune disorder; angiogenic disorder;
 KW gene therapy; cytostatic; neuroprotective; gene; ss.

OS Homo sapiens.

PN WO200153486-A1.

26-JUL-2001.
PD

11-FEB-2000: 2000WO-US003565.

XX
08-MAR-1999. 99W0-11S005028
PP

PR 11-MAR-1999; 99US-0123972P.
PR 11 MAY 1999; 99US 0123450D

PR 02-JUN-1999; 99WO-US012252.

PR 22-JUN-1999; 99US-0140653P.

PR 26-JUL-1999; 99US-0145698P.

FR 28-007-1999; 99US-0149395P.
PR 17-AUG-1999; 99US-0149395P.

PR 31-AUG-1999; 99US-0151689F.
PR 01-SEP-1999: 99WO-US020111.

PR 15-SEP-1999; 99WO-US021090.
PR 30-NOV-1999; 99WO-US028313

PR	01-DEC-1999;	99WO-US028301.
PR	01-DEC-1999;	99WO-US028301.
PR	01-DEC-1999;	99WO-US028301.

PR 05-JAN-2000; 2000WO-US000219.

PA (GETH) GENENTECH INC.

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;

PI Marsters SA, Pan J, Pizzi RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK. Wood WI:

XX
WPT: 2002-205567/26

DR P-PSDB; AAU86129.

PT Thirty five nucleic acids encoding PRO polypeptides, useful for t
PT benign or malignant tumors, leukemias and lymphoid malignancies,
PT inflammatory, angiogenic and immunologic disorders.

vention relates to the isolation of novel human PRO and the polynucleotide sequences encoding them. The PRO agonists, antagonists or anti-PRO antibodies are useful for gn or malignant tumors (e.g. renal, kidney, bladder, leukemias and lymphoid malignancies, other disorders such as glial, astrocytal, hypothalamic, glandular, macrophagal, metastasocelic disorders, inflammatory, immune and angiogenic). The polynucleotide sequences are also useful in gene therapy. The 0288 encode for the human PRO polypeptides of the invention BP: 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

53.8%; Score 628.6; DB 6; Length 1353;
arity 76.2%; Pred.No.1.2e-159;
onservative 0; Mismatches 219; Indels 73; Gaps 10;

TGTGACCTGGGCCTGGCGTGGCCCTGCCCTTGCCCTCCTGCTCGTGGTCGTCAGCCCTG 61
TTGGCCCTGGGCCTGGCGTGGCCCTGCCCTGGCCCTCCTGCTGGCCGTGGTCAGTTTG 189
TAGCTGGCAACCGCTGTCTGCCAGGAGCCCTTCTCAGSAGAGCTGCACAGAGGAC 121
AGCCGGGCATCGCTGTCGCCCCAGGAGCCCTGCCAGGAGAGCTGGTGGCAGAGGAG 249
GGGAGCCCCCTGAATGAACTCCCAGACAGAGGAAGCCAGGATGTGTACTCTTC 181
TAGAACCCGCTCGAATGAACTGCCAGACAGAGAAGAAGCCAGGATCTCGGCCCTTC 309
AACAACATAGTCGGCCCTCGAAGAAGTGCTCTAAAGCCGGAAGCGCGCCCTCGC 241
AACCGCATAGTTGCGCCCTCGAAGTGCACCTTAAGGCCGGAACACGGGCTCGA 369
GCTATTGCAGCCCATPATGAGGTTTCATCTCGCCAGGACAGGATGGAGCACAGCA 301
GCGATCGAGCCCATTAAGAAGTTTATCCACAGCTTGGACAGGACGGAGCGCAGGCA 429
TGTGATGGAGACGATGAGTGGCTGGGAAGAGACCAAATCAACAGCTCCAGCCCTCG 361
TGTGAGCGGACACGTAGTGGCTGGAGGAGACAGGAATCAACAGCTCCAGCCCTCG 489
TACAGACCGCCAGATTCGGGAAATTTACAGTCATCAGGCTCGGGCTCTACTACTCTGAC 421
TACAACCGCCAGATCGGGAGTTTATGTCACCGGGCTGGGCTCTACTACTCTGAC 549
CAGGTGTCATTTGATGAGGAAAAGGCTGTCTACTGAAGCTGGACTTCTGCTGAAC 481
CAGGTGTCATTTGATGAGGGAAGGCTGTCTACTGAAGCTGGACTTCTGCTGGAT 609
TGTGTCGGCCCTGGCTGCGTGGGAAGAAATCTCAGCCACAGCAGCAAGCTCTCCTGG 541
TGTGTCGGCCCTGGCTGCGTGGAGAAATCTCAGCCACTCGGCGAGTCCCTCGGG 669
CAGCTCCGTTTGTGCAGAGTGTCTGGGCTGTGCGCTGTGGCCCGCAGGCTCTTCCCTT 601
CAGCTCCGCTCTGACCAGTGTCTGGCTGTGGGCTGTGCGCTGTGGCCCGCAGGCTCTTCCCTT 729
TATCCGACCCCTCCCTGGGCTCATCTTAAAGGTGCCCCCTTCTTAACCTACTTTTGA 661
TATCCGACCCCTCCCTGGGCCATCTCAAGSGTCGCCCTTCTCTCACTACTTGGGA 789
TTTTCAAGTTCACTGAGGGGCTTGCTCTCCAGATTCTTAAACTTTCCCTGGCTCC 721
TTTCCAGGTTCACTGAGGGGCTTGCTCTCCCGCAGTCTGCCAGGCTGCGGGCTCC 849
IAGCATCACACACACTCCCTACCCCAACCCCACTCCTCAGCCCCCTC-GCTGCTCTT 780
---CCTGCACAGCTCTGTGGGACCCGGTCCCTCTGCCCCACCTCTCAGCGCTCTTT 904
CCAGTCTGTCTCTCC--TCAAGGACGACGAGCTTGTTCACATGTTTCCATTCC- 837

905	GCTCCAGACCTGCCCTCCCTCTAGAGGCTGCTGGGCTCTGTTTCAAGTGTGTTTC	Db
838	-----ACAGACGTATCTTGTCTTCTTTAACATCCCATCCACACAACTATCC	QY
965	ACATAAATACAGTATTCCTTATCTTAACTCTCCCTCCACCGCCCACTCTCC	Db
892	ACTAGTCTCCCAAGGCCCTTAC-----TTATCCTGACTCCCCCACC	QY
1025	ACTAGCTCCCAATCCCTGACCCCTTTGAGGCCCTCAGTGATCTCGACTCCGCCCT	Db
937	CACCGCACCACTGTTTATTATGACTTTGTGCAC-----	QY
1085	CAGACCCCCAGGTCATTGTGTTTCACTGTACTCTGTGGGCAAGATGGGTCCAGAA	Db
969	-----CAGGCACTGAGATGGGCTGGACCTGGTGCGAGGAAGCCAGAGAACTCTGGG	QY
1145	CAC TTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAAGGAGGCAAGACACTGGG	Db
1074	GCCAGAGTTCCTCCAACTGTCAGGGGGAGAGCTGGGGACAAGCTCCTCCCTCGA-	QY
1205	GCCAGGAGTTCCTCCAAATGTGAGGGGGGAGA-AACAAGACAAGCTCTCTCCCTTGAG	Db
1080	CTGTGGATTTTGAAA--AGATACATTTTTTATTATTATTGTCACAAAATGT---	QY
1264	CTGTGGATTTTAAACACAGATATTATTTTATTATTATTGTCACAAAATGTGTA	Db
1135	GGATTTAAAGAGAATAAATCATGA 1159	QY
1324	GGATTTAAATAGAAATAGTCAATA 1348	Db

RESULT 6
AA56000
ID AAX56000 standard: DNA: 1421 BP

XX	AAX56000;	
XX	AC	
XX	DT	
XX	15-JUL-1999	(first entry)
XX	DE	
XX	Human tumour necrosis factor Apo-3	ligand polynucleotide sequence
XX	XX	
XX	Human; tumour necrosis factor; Apo-3	ligand; lymphotoxin; apopto
KW	NF-kappaB-dependent transcription;	JNK/SAPK-dependent response;
KW	ss.	
XX	XX	
XX	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
FH	CDS	92..841
FT		/*tag= a
FT		/product= "Apo-3 ligand"
XX	XX	
PN	WO9919490-A1.	
XX	XX	
PD	22-APR-1999.	
XX	XX	
PF	09-OCT-1998;	98WO-US021407.
XX	XX	
XX	10-OCT-1997;	97US-0062037P.
PR	17-DEC-1997;	97US-0069862P.
XX	XX	
PA	(GETH) GENENTECH INC.	
XX	XX	
PI	Ashkenazi AJ, Marsters SA, Pitti R;	
XX	XX	
DR	WPI; 1999-287982/24.	
DR	P-PSDB; AAY09369.	
XX	XX	
PT	New human Apo3- ligand (a tumor necrosis factor) homologue.	
XX	XX	
PS	Claim 18; Fig 1; 74pp; English.	
XX	XX	
CC	The present sequence encodes a human tumour necrosis factor (TNF	

Conservative 0; Mismatches 220; Indels 73; Gaps 10;

1024 GCCAGAGTTCCCAACTGTGTAGGGGGAAGAGCTGGGGAACAAGCTCCTCCCTGGA--
 1165 GCCAGGAGTTCCCAAAATGTGAGGGGCGGAGA-AACAAGACAAGCTCCTCCCTTGAG#
 1080 CCTGTGGATTTTGAAA--AGATACTATTTTATTATTATTGACAAAATCT---1
 1224 CCTGTGGATTTTAAACAGATATTTATTATTATTATTATTGACAAAATGTTGA#
 1135 GGATATTAAAGAGATAAATCA 1156
 1284 GGATATTAAATAGATAAGTCA 1305

RESULT 8
 ACC57901
 ID ACC57901 standard; cDNA; 1306 BP.
 XX ACC57901;
 AC ACC57901;
 DT 11-AUG-2003 (first entry)
 XX Human TWEAK coding sequence.
 DE Human; TWEAK; tumour necrosis factor; ligand; cytostatic;
 KW immunomodulator; osteopathic; gene; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 18..767
 FT /*tag= a
 FT /product= "Human TWEAK"
 XX WO2003040307-A2.
 PN 15-MAY-2003.
 XX 25-JUL-2002; 2002WO-US023782.
 XX 27-JUL-2001; 2001US-0307838P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Hilbert DH, Rosen CA;
 PI WPI; 2003-430659/40.
 DR P-PSDB; ABR42315.
 XX New heteromultimeric complex having a first polypeptide member o
 PT tumor necrosis factor (TNF) ligand family, and a second differen
 PT of TNF ligand family, useful for treating cancer, osteoporosis o
 PT autoimmune disease.
 XX Disclosure; Page 367-368; 388pp; English.
 XX The present sequence is that of a polynucleotide encoding human '
 CC The invention relates to compositions comprising heterotrimeric '
 CC of tumour necrosis factor (TNF) ligand family members, and their
 CC the detection, prevention and treatment of disease. In one embod
 CC the heterotrimeric complex comprises full-length or extracellular
 CC portions of TWEAK and full-length or extracellular portions of o
 CC ligand family members, preferably VEG1 or VEG1-SV. The heterotri
 CC complexes of the invention are useful for treating an autoimmune
 CC cancer or osteoporosis, and particularly for inhibiting cancer c
 CC proliferation, increasing B cell proliferation, or inducing apop
 CC T cells
 XX SQ Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match 53.4%; Score 624; DB 7; Length 1306;
 Best Local Similarity 76.0%; Pred. No. 2.1e-158;
 Matches 929; Conservative 0; Mismatches 220; Indels 73;

06:25:10 2004

Conservative 0; Mismatches 220; Indels 73; Gaps 10;

1024 GCCAGAGTTCCCAACTGTGTAGGGGGAAGAGCTGGGGAACAAGCTCCTCCCTGGA--
 1165 GCCAGGAGTTCCCAAAATGTGAGGGGCGGAGA-AACAAGACAAGCTCCTCCCTTGAG#
 1080 CCTGTGGATTTTGAAA--AGATACTATTTTATTATTATTGACAAAATCT---1
 1224 CCTGTGGATTTTAAACAGATATTTATTATTATTATTATTGACAAAATGTTGA#
 1135 GGATATTAAAGAGATAAATCA 1156
 1284 GGATATTAAATAGATAAGTCA 1305

RESULT 8
 ACC57901
 ID ACC57901 standard; cDNA; 1306 BP.
 XX ACC57901;
 AC ACC57901;
 DT 11-AUG-2003 (first entry)
 XX Human TWEAK coding sequence.
 DE Human; TWEAK; tumour necrosis factor; ligand; cytostatic;
 KW immunomodulator; osteopathic; gene; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 18..767
 FT /*tag= a
 FT /product= "Human TWEAK"
 XX WO2003040307-A2.
 PN 15-MAY-2003.
 XX 25-JUL-2002; 2002WO-US023782.
 XX 27-JUL-2001; 2001US-0307838P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Hilbert DH, Rosen CA;
 PI WPI; 2003-430659/40.
 DR P-PSDB; ABR42315.
 XX New heteromultimeric complex having a first polypeptide member o
 PT tumor necrosis factor (TNF) ligand family, and a second differen
 PT of TNF ligand family, useful for treating cancer, osteoporosis o
 PT autoimmune disease.
 XX Disclosure; Page 367-368; 388pp; English.
 XX The present sequence is that of a polynucleotide encoding human '
 CC The invention relates to compositions comprising heterotrimeric '
 CC of tumour necrosis factor (TNF) ligand family members, and their
 CC the detection, prevention and treatment of disease. In one embod
 CC the heterotrimeric complex comprises full-length or extracellular
 CC portions of TWEAK and full-length or extracellular portions of o
 CC ligand family members, preferably VEG1 or VEG1-SV. The heterotri
 CC complexes of the invention are useful for treating an autoimmune
 CC cancer or osteoporosis, and particularly for inhibiting cancer c
 CC proliferation, increasing B cell proliferation, or inducing apop
 CC T cells
 XX SQ Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match 53.4%; Score 624; DB 7; Length 1306;
 Best Local Similarity 76.0%; Pred. No. 2.1e-158;
 Matches 929; Conservative 0; Mismatches 220; Indels 73;

1024 GCAGAGTTCCTCCAACTGTGAGGGGAGAGCTGGGGACAAGCTCTCCTCGA--
1165 GCAGAGTTCCTCCAACTGTGAGGGGAGAGCTGGGGACAAGCTCTCCTCGA.
1080 CTGTGCTGATTTGAAA--AGATACTATTTTATTATTATTTGACAAAATGT--T.
1224 CTGTGCTGATTTTAAACAGATATTTATTTTATTATTATTGACAAAATGTTAT.
1135 GGATATTAAAGAGATAAATCA 1156
1284 GGATATTAAAGATAAATCA 1305

RESULT 9
ADC35205
ID ADC35205 standard; cDNA; 1306 BP.
XX ADC35205;
AC ADC35205;
XX XX
DT 18-DEC-2003 (first entry)
XX
DE Human cDNA encoding TNF ligand family member #12.
XX
XX ss; gene; human; tumour necrosis factor; TNF ligand; endokine alp
KW excessive bone resorption disorder; osteoporosis; Paget's disease
KW arterial calcification.
XX
XX Homo sapiens.
XX
XX US2003100074-A1.
XX
XX 29-MAY-2003.
XX
XX 15-AUG-2002; 2002US-00218547.
XX
XX 16-AUG-2001; 2001US-0312542P.
PR 30-OCT-2001; 2001US-0330761P.
XX
XX (YUGG/) YU G.
PA (NIJ/) NI J.
PA (ROSE/) ROSEN C A.
PA (NARD/) NARDELLI B.
XX
XX Yu G, Ni J, Rosen CA, Nardelli B;
XX
XX WPI: 2003-696072/66.
DR P-PSDB; ADC35205.
XX
XX New Endokine alpha gene useful for preparing a composition for tr
PT disease associated with excessive or insufficient bone resorption
PT osteoporosis, Paget's disease or arterial calcification.
XX
XX Disclosure; SEQ ID NO 23; 145pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encodi
CC tumour necrosis factor family ligand. A composition comprising th
CC isolated antibody or its fragment is used for treating an individ
CC need of decreased level of endokine alpha activity. The endokine
CC polypeptide present in a heterotrimeric complex is used for treat
CC individual having a disorder associated with excessive bone resor
CC e.g. osteoporosis, Paget's disease or arterial calcification. Tre
CC individual having a disorder associated with insufficient bone re
CC comprises administering an endokine alpha antagonist, which is th
CC antibody that binds specifically to endokine alpha polypeptide. T
CC present sequence represents a cDNA encoding a tumour necrosis fac
CC family ligand.
XX
XX Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;
SQ

Query Match 53.4%; Score 624; DB 9; Length 1306;
Best Local Similarity 76.0%; Pred. No. 2.le-158;
Matches 929; Conservative 0; Mismatches 220; Indels 73; G

51.2%; Score 597.8; DB 4; Length 1236;
arity 75.2%; Pred. No. 2.5e-152;
onservative 0; Mismatches 222; Indels 68; Gaps 8;
TGAGCTGGGGCTGGCGCTGGCCCTGCCTTGGCCTCTCTGCTGGTCTGGTCTGAGCCTG 61
TGCGCTGGGCTGGCGCTGGCCTGCCTTGGCCTCTCTGCGCCTCTGCTGGCGTGGTCA GTTTG 132
GCTGGGCAA CGCTGTCTGCCAGAGAGCCTTCTCAGGAGGAGCTGACAGCAGAGAC 121
GCCGGGATCGCTGTCCGCCAGAGAGCCTGCCCAGAGGAGCTGGTGGCAGAGAG 192
GGGAGCCCCCTGAACTGAATCCCAGACAGAGGAAAGCCAGATGTGTACTCTTTC 181
AGGACCCCTCGAACTGAATCCCAGACAGAAAGAACCGAGATCTCTGCGCCTTTC 252
AACAACCTAGTCCGCGCTCGAAGAGTGTCTTAAAGCGCGAAGCGCGCCTCGC 241
LACGACTAGTTCCGCGCTCGAAGAGTGCACCTTAAAGGCGCGAAACAGCGGCTCGA 312
CTATTTCAGAGCCATTATCAGGTTTCATCTCCGCGAGGACAGATCGAGCAACAAGCA 301
CGATTCGAGAGCCATTATGAAGTTTATCCAGACCTGGAACAGACGAGCGCAGGCA 372
TGGATGGGACAGTGAAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCTG 361
TGGAGCGGACAGTGAAGTGGCTGGGAGGAAGCAGAAATCAAAGCTCCAGCCCTCTG 432
TAGGACCGCAGATTGGGGAATTTACGTATCAGGGCTGGGCTCTACTACCTGTAC 421
TACAACCGCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTCTACTACTGTATC 492
TAGGTGCACTTTGATCAGGAAAGGCTGTCTACTCAAGCTGAGCTTGTGTGTGAAC 481
TAGGTGCACTTTGATCAGGGGAAGGCTGTCTACTGAAGCTGAGCTTGTGTGTGAT 552
TGTCTGGCCGTGGCTGCTGGGAAGAAATTTCTAGCCACAGCAGCAAGCTCTCTGGG 541
TGTCTGGCCGTGGCTGCTGGGAAGAAATTTCTAGCCACCTGCGCGAGTTTCCCTCGG 612
TAGCTCGGTTTGTGCAGGTTCTGGGCTGTTCGCGCTGGCGCCAGGGTCTTCCCTT 601
TAGCTCGGCTCTGCCAGGTTCTGGGCTGTTCGCGCTGGCGCCAGGGTCTTCCCTT 672
ATCCGACCTCCCTGGGCCCCTCACTCAAGGCTGCCCCCTTCTCCTACTTCTGGA 661
ATCCGACCTCCCTGGGCCCCTCACTCAAGGCTGCCCCCTTCTCCTACTTCTGGA 732
TTTCAAGTTCACTGAGGGGCTTGTCTCTCCAGATTCCTTAAACTTTTCCCTGGCTCC 721
TTCCAGGTTCACTGAGGGGCTTGGTCTCCCGAGTGTCTCCAGGCTGTCGCGAGGCTCC 792
AGCATCACCACACCTCCCTACTACCCACACCCCACTCTCCACCCCTC-GCTGCTCCTT 780
--CCTGCAAGCTCTCTGGGCAACCGGCTCCCTCTGCCCCACCTCAGCGCTCTTT 847
CCAGTCTGTCTCTCC-TCAAAGGCGAGCAGAGCTTGTGTCAATGTTTCCATTC- 837
CCAGACTGCCCCCTCTCTAGAGGCTGCGCTGGGCTTGTTCAGGTGTTTTCATCCC 907
---ACAGAGTATCTTGTCTTCTTAAATCCATCCACCAACAATATCCACCTC 891
TAAATACAGTATTCCTCACTTTATCTTACAAACCCCAACCGGCCATCTTCCACCTC 967
AGCTCCCAAGCCCCCTAC-----TTATCCCTGACTCCCCCAACCCCACT 936
AGCTCCCAATCCCTGACCTTTGAGGCCCCCACTGATCTCGACTCCCCCTGGCCA 1027
CCGACACAGGTTTATTTGACTTTGTGAC-----968
TACCCCAAGGCTTGTGTTTCACTGTACTGTGGGCAAGATGGTCCAGAGACCC 1087

QY	969	-----CAGGCACTGAGATGGCTGGACCTGGTGGCAGGAGCCAGAGCACTGGG	
Db	1088	CACCTTCAGGCACTAAGAGGGGGCTGGACCTGGCGGCGAGGAGCCAAAGAGACTGGG	
QY	1024	GCCAGAGCTCCCAACTGTGAGGGGGAAGAGCTGGGGCAAGCTCCTCCCTGGGA--	
Db	1148	GCCAGAGCTCCCAACTGTGAGGGGGAAGAGCTGGGGCAAGCTCCTCCCTGGGA--	
QY	1080	CCTGTGGATTTTGAAGAATCTATTTTT 1108	
Db	1207	CCTGTGGATTTTAAACAGATATTTTT 1235	
RESULT 14			
AAAX23424			
ID	AAAX23424	standard; DNA; 1030 BP.	
AC	AAAX23424;		
XX	19-JUN-1999	(first entry)	
DT			
XX			
XX			
DE	Human TNRL3 DNA.		
XX			
KW	Tumour necrosis factor receptor; signal transducer molecule; TNF		
KW	Developmental abnormality; Gestational abnormality; prostate c		
KW	AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy;		
KW	cytoplasmic domain; immunogen; antibody preparation; breast carc		
KW	apoptosis; human; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
PH	1..627		
FT	/*tag= a		
FT	/product= "TNRL3"		
FT			
XX			
PN	WO9911791-A2.		
XX			
PD	11-MAR-1999.		
XX			
XX	04-SEP-1998; 98WO-US018393.		
PF			
XX			
PR	05-SEP-1997; 97US-00924634.		
XX			
PA	(UNIW) UNIV WASHINGTON.		
XX			
PI	Chaudhary PM;		
XX			
DR	WPI; 1999-205191/17.		
DR	P-PSDB; AAW93590.		
PT	New Tumor Necrosis Factor family receptor polypeptides and ligat		
PT	useful for diagnosis and treatment of prostate cancer and develo		
PT	or gestational abnormalities.		
XX			
PS	Example VII; Fig 13A; 156pp; English.		
XX			
CC	This invention describes isolated Tumor Necrosis Factor (TNF) fa		
CC	receptor polypeptides: APO4, APO6, APO8 and APO9 or their active		
CC	fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and 1		
CC	their active fragments. APO4 is useful for diagnosing prostate c		
CC	determining levels of APO4 in an individual. Prostate cancer can		
CC	treated using APO4 selective binding agents linked to a therapeu		
CC	moly. APO4 polypeptides are also useful for identifying select		
CC	binding agents, useful in diagnosis/treatment of disease by bind		
CC	agents to the polypeptide/active fragment which is extracellular		
CC	expressed on the cell surface. The binding is preferably perform		
CC	vivo. APO4 polypeptides/ active fragments are also useful for sc		
CC	for agonists and antagonists by binding and observing the change		
CC	activity. Effective pharmacological agents useful in diagnosis c		
CC	treatment of disease are also identified using APO4 polypeptides		
CC	fragments and APO4 signal transducer molecules that specifically		
CC	with a cytoplasmic domain of APO4 and detecting a change in leve		

06:25:10 2004

us-09-245-198a-1.rng

42.7%; Score 498.8; DB 4; Length 898;
arity 87.0%; Pred No 1.4e-124;
onservative 0; Mismatches 82; Indels 0; Gaps 0;

QY TGGGAGCTGGGCAACGCTGTCTGCCAGAGGACCTTCTCAGGAGAGCTGACAGCA 115
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
TGGGAGCCGGGCATCGCTGTCCGCCAGGAGCCTGCCAGGAGAGCTGGTGGCA 309
QY AACGCCGGAGCCCCCTGAATCTGATCCCCACAGAGGAAAGCCAGGATGTGTA 175
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
AAGCACAGGACCCGTCGGAACCTGAATCCCCACAGAGGAAAGCCAGGATCCTGG 369
QY TCTTGGAAACAATAGTCCGGCCTCGAAGAAAGTCTCTAAAGCCGGHAAGCGCGG 235
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
TCTTGAACCGACTAGTTCGGCCTCGCAGAGTGACACCTAAAGCCGGAAACACGG 429
QY GCGGAGCTATTCCAGCCCATTTATGAGGTTTCATCTCGGCCAGGACAGGATGGACA 295
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
GAAGAGCGATCGACGCCCATTTATGAAGTTTCATCCACGACCTGGACAGGACGAGCG 489
QY CAGGTGTGGATGGGACAGTGTGCTGGAGAGACCAAAATCAACAGCTCCAGC 355
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CAGGTGTGACCGGACAGTGTGCTGGAGAGACCAAGATCAACAGCTCCAGC 549
QY TCGGCTACGACCGCCAGATTTGGGGAATTTACAGTCATCAGGCTGGGCTCTACTAC 415
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
TGGCCTACAAACCGCCAGATCGGGAGTTTATAGTCACCGGGCTGGGCTCTACTAC 609
QY TCTGTCAAGTGACATTTGATGAGGGAAGGCTGTCTACTGAAGCTGACTTCTG 475
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
TCTGTCAAGTGACATTTGATGAGGGAAGGCTGTCTACTGAAGCTGACTTCTG 669
QY ACGGTGTGTGGCCCTGGCTGCCCTGGAGAAATTTCTCAGCCACAGCAGCAAGCTCT 535
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
ATGGTGTGTGGCCCTGGCTGCCCTGGAGAAATTTCTCAGCCACTGGGCCAGTTCC 729
QY GGGCCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTTCGGCTGGGCCAGGGTCT 595
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
GGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTTCGGCTGGGCCAGGGTCC 789
QY TTCGGATCCGACCCCTCCGCTGGGCTCATCTTAAGGCTGCCCCCTTCTTAACCTAC 655
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
TTCGGATCCGACCCCTCCGCTGGGCCCATCTCAAGGCTGCCCCCTTCTTAACCTAC 849
QY GACTCTTTCAAGTTCACTGAGGGGCC 685
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
GACTCTTCCAGTTCACTGAGGGGCC 879

April 7, 2004, 21:32:17
1 secs

GenCore version 5.1.6
copyright (c) 1993 - 2004 Compugen Ltd.

n search, using sw model

il 7, 2004, 17:38:07 ; Search time 9.28291 Seconds
(without alignments)
1262.081 Million cell updates/sec

09-245-198A-2

LSLGLALACIGLLVWSL.....PWAHLKAAPLTFLVGLFQVH 225

SUM62

top 10.0 , Gapext 0.5

681 seqs, 52070155 residues

s satisfying chosen parameters: 141681

h: 0

h: 2000000000

imum Match 0%

ximum Match 100%

sting first 45 summaries

visProt_42:*

the number of results predicted by chance to have a
score than or equal to the score of the result being printed,
and by analysis of the total score distribution.

SUMMARIES

seq	ch	Length	DB	ID	Description
0.0	225	1	TN12	MOUSE	O54907 mus musculus
7.8	249	1	TN12	HUMAN	O43508 homo sapien
9.3	272	1	TN15	CHICK	O918d8 gallus gall
3.8	316	1	TN11	MOUSE	O35235 m tumor nec
7.8	260	1	TN15	CANFA	O97626 canis famil
7.8	318	1	TN11	RAT	O95922 x tumor nec
7.6	532	1	PBEN	HUMAN	P10696 homo sapien
7.6	3664	1	MINT	HUMAN	O96t58 homo sapien
7.5	244	1	TNFC	HUMAN	O96643 homo sapien
7.5	244	1	TNFC	PANTR	O86227 pan troglod
7.5	261	1	TN15	AOTTR	O9bdm3 aotus trivi
7.5	261	1	TN15	CALJA	O9bdm3 callithrix
7.4	240	1	TN14	HUMAN	O43557 homo sapien
7.4	1237	1	B3A2	MOUSE	P13808 mus musculus
7.4	261	1	TN15	HUMAN	P29965 homo sapien
7.4	261	1	TN15	MACMU	O9bdc7 macaca mula
7.3	240	1	TN15	MAGNE	O9bdm7 macaca neme
7.3	1237	1	B3A2	RABIT	P48746 oryctolagus
7.2	278	1	TN16	RAT	P36940 rattus norv
7.2	241	1	TN13	MOUSE	O9d777 mus musculus
7.2	250	1	TN13	MACEU	O9xt47 macropus eu
7.1	400	1	TRPB	CHROMO	Q7nud8 chromobacte
7.1	2779	1	LVA	BROME	O8msl1 drosophila
7.1	535	1	PB1	HUMAN	P05187 homo sapien
6.9	246	1	CLQC	MOUSE	O02105 mus musculus
6.9	1234	1	B3A2	RAT	P23347 rattus norv
6.9	1465	1	DPO3	STRMU	O8dwe0 streptococc
6.8	197	1	TN16	RABIT	P10154 oryctolagus
6.8	300	1	NTH1	MOUSE	O35980 mus musculus
6.8	920	1	PARC	SYN13	P73077 synecocyst
6.8	1584	1	U104	CABEL	P23678 caenorhabdi
6.8	2468	1	MAPB	HUMAN	P46821 homo sapien
6.8	260	1	TN15	FELCA	O97605 felis silve

34	78.5	6.8	480	1	KCG2_RAT	Q9qru3 ratt
35	78.5	6.8	1164	1	PHYD_ARATH	P42497 arab:
36	78	6.7	285	1	T13B_HUMAN	Q9Y275 homo
37	78	6.7	495	1	GATB_METAC	Q8thj0 meth:
38	78	6.7	495	1	GATB_METMA	Q8px10 meth:
39	78	6.7	763	1	APP2_HUMAN	Q06481 homo
40	77.5	6.7	201	1	TNFB_WACEU	Q9xt48 macr:
41	77.5	6.7	261	1	TNFB_PIG	Q95mq5 sus:
42	77	6.6	279	1	TNFB_MOUSE	P41047 mus:
43	77	6.6	817	1	NAH1_BOVIN	Q28036 bos:
44	76.5	6.6	214	1	SMP_ECOLI	P18838 esch:
45	76.5	6.6	788	1	NASP_HUMAN	P49321 homo

ALIGNMENTS

RESULT 1
TN12_MOUSE
ID TN12_MOUSE STANDARD; PRT; 225 AA.
AC O54907; Q9CTP2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE tumor necrosis factor ligand superfamily member 12 (TNF-related w
inducer of apoptosis) (TWEAK) (fragment).
GN TNFSF12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OC NCBI_TaxID=10090;
CX [1]
RN TNFSF12
RP TISSUE=Peritoneal macrophage;
RX MEDLINE=98070415; PubMed=9405449;
RA Chicopeptiche Y.; Bourdon P.R.; Xu H.; Hsu Y.-M.; Scott H.;
RA Hession C.; Garcia I.; Browning J.L.;
RT "TWEAK, a new secreted ligand in the tumor necrosis factor family
weakly induces apoptosis."
RL J. Biol. Chem. 272:32401-32410(1997).
[2]
SEQUENCE OF 83-225 FROM N.A.
STRAIN=C57BL/6J; TISSUE=Retina;
MEDLINE=21085660; PubMed=11217851;
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush
Schrml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washic
Sakai K., Okido T., Furuno M., Aono H., Baldarelli J., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmir
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -!- FUNCTION: Binds to FN14 and possibly also to TNFSF12/APO3. A
inducer of apoptosis in some cell types. Promotes angiogenesis
the proliferation of endothelial cells. Mediates NF-KappaB
activation (by similarity).
CC -!- SUBUNIT: Homotrimer (Potential).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted
similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- PTM: The soluble form is produced from the membrane form by

ic processing (By similarity).
Y: Belongs to the tumor necrosis factor family.

OT entry is copyright. It is produced through a collaboration Swiss Institute of Bioinformatics and the EMBL outstation - Bioinformatics Institute. There are no restrictions on its profit institutions as long as its content is in no way this statement is not removed. Usage by and for commercial aires a license agreement (See <http://www.isb-sib.ch/announce/mail-to:license@isb-sib.ch>).

10; AAC53517.1; -
19; BAB32249.1; -
1259; Tnfef12.
1006052; TNF family.
1008983; TNF-like.
; TNF; 1.
17; TNF; 1.
1251; TNF 1; FALSE_NEG.
1049; TNF 2; 1.
Hogensis; Apoptosis; Transmembrane; Glycoprotein;
1; 1
1 225 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
MEMBER 12, MEMBRANE FORM.
70 225 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
MEMBER 12, SECRETED FORM (BY SIMILARITY).
1 21 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
22 225 EXTRACELLULAR (POTENTIAL).
69 70 CLEAVAGE (BY SIMILARITY).
15 115 N-LINKED (GLCNAC...) (POTENTIAL).
15 AA; 24781 MW; 90C412C0480659B CRC64;
100.0%; Score 1162; DB 1; Length 225;
arity 100.0%; Pred. NO. 2.5e-95;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GLALACGLLVVSLGSGWATLSAQEPQELTAEDEPPELNPQTESQDVVPF 60
GLALACGLLVVSLGSGWATLSAQEPQELTAEDEPPELNPQTESQDVVPF 60
VPRPSAPKGRARRAIAAHYVHPRPGDGAQAGVDGTVSGWEETKINSSPL 120
VPRPSAPKGRARRAIAAHYVHPRPGDGAQAGVDGTVSGWEETKINSSPL 120
IQIGFTVIRAGLYLYCQVHFDEGKAVYKLDLLVNGVLALRCLEFSAATASSPG 180
IQIGFTVIRAGLYLYCQVHFDEGKAVYKLDLLVNGVLALRCLEFSAATASSPG 180
ILCQVSGLLPLRPGSSLRITLPAWHLKAAPFLTYGLFQVH 225
ILCQVSGLLPLRPGSSLRITLPAWHLKAAPFLTYGLFQVH 225

STANDARD; PRT; 249 AA.
27;
(Rel. 41, Created)
(Rel. 41, Last sequence update)
(Rel. 42, Last annotation update)
is factor ligand superfamily member 12 (TNF-related weak
poptosis) (TWEAK) (APO3 ligand).
PO3L OR DR3LG.
(Human).
etazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Primates; Catarrhini; Hominidae; Homo.
506;
M N.A., AND N-TERMINUS OF SOLUBLE FORM.
liver, and Tonsil;

RA MEDLINE=98070415; PubMed=9405449;
RA Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RA Hession C., Garcia I., Browning J.L.;
RT "TWEAK, a new secreted ligand in the tumor necrosis factor family
RT weakly induces apoptosis.";
RL J. Biol. Chem. 272:32401-32410(1997).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=98228355; PubMed=9560343;
RA Masters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,
RA Ashkenazi A.;
RT "Identification of a ligand for the death-domain-containing recep
RT Apo3.";
RL Curr. Biol. 8:525-528(1998).
[3]
RN SEQUENCE FROM N.A.
RP TISSUE=Tonsil;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.I.
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.I
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RN FUNCTION.
RP MEDLINE=99185061; PubMed=10085077;
RA Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S
RT "TWEAK induces angiogenesis and proliferation of endothelial cel
RL J. Biol. Chem. 274:8455-8459(1999).
CC -!- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/AP03. I
CC inducer of apoptosis in some cell types. Mediates NF-kappaB
CC activation. May promote angiogenesis and the proliferation o
CC endothelial cells.
CC -!- SUBUNIT: Homotrimer (Potential).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted.
CC -!- TISSUE SPECIFICITY: Highly expressed in adult heart, pancrea
CC skeletal muscle, brain, colon, small intestine, lung, ovary,
CC prostate, spleen, lymph node, appendix and peripheral blood
CC lymphocytes. Low expression in kidney, testis, liver, placen
CC thymus and bone marrow. Also detected in fetal kidney, liver
CC lung and brain.
CC -!- PTM: The soluble form derives from the membrane form
CC by proteolytic processing.
CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift in position 125.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a col
CC between the Swiss Institute of Bioinformatics and the EMBL ou
CC the European Bioinformatics Institute. There are no restrictio
CC use by non-profit institutions as long as its content is i
CC modified and this statement is not removed. Usage by and for
CC entries requires a license agreement (See <http://www.isb-sib.ch>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF030099; AAC51923.1; -
CC EMBL; AF055872; AAC39724.1; -
DR
DR

```

; AAH19047.1; ALT_FRAME.
; 927; TNFSF12.
; C: integral to plasma membrane; TAS.
; F: receptor binding; TAS.
; P: induction of apoptosis; TAS.
; S: signal transduction; TAS.
; 06052; TNF family.
; 08983; TNF-like.
; TNF; 1.
; TNF; 1.
; 51; TNF_1; FALSE_NEG.
; 149; TNF_2; 1.
; 0genesis; Apoptosis; Transmembrane; Glycoprotein;
1 249 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
MEMBER 12, MEMBRANE FORM.
14 249 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
MEMBER 12, SECRETED FORM.
1 21 CYTOPLASMIC (POTENTIAL).
22 42 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
43 249 EXTRACELLULAR (POTENTIAL).
33 94 CLEAVAGE.
39 139 N-LINKED (GLCNAC. . .).
3 AA; 27216 MW; E660843361C28EBA CRC64;
87.8%; Score 1020; DB 1; Length 249;
arity 88.8%; Pred. No. 9.4e-83;
conservative 9; Mismatches 16; Indels 0; Gaps 0;
LALACGLLVVSLGSWATLSAQPSOELTAEDRRPELNPQTEESQDVVPFL 61
LALACGLLVVSLGSRASLSQAEPAQELVAEDQDPSLNPQTEESQDPAPFL 85
RPRRSAPKGRKARRAIAAHYEVHPRPGDGAQGVDTGVSQWEERINSSSPLR 145
RPRRSAPKGRKARRAIAAHYEVHPRPGDGAQGVDTGVSQWEERINSSSPLR 145
IGFTVIRAGLYLYCQVHFDEGKAVYKLDLLVGVLAIRCLLEFSATASSPGP 181
IGFTVIRAGLYLYCQVHFDEGKAVYKLDLLVGVLAIRCLLEFSATASSPGP 181
IGFTVIRAGLYLYCQVHFDEGKAVYKLDLLVGVLAIRCLLEFSATASSPGP 205
QVSGLLALRPGSSLRITLPWHLKAAPFLTYFGLFQVH 225
QVSGLLALRPGSSLRITLPWHLKAAPFLTYFGLFQVH 249
QVSGLLALRPGSSLRITLPWHLKAAPFLTYFGLFQVH 249
STANDARD; PRT; 272 AA.
Rel. 41, Created
Rel. 43, Last sequence update
Rel. 43, Last annotation update
s factor ligand superfamily member 5 (CD40 ligand) (CD40-
tein)
OLG OR CD40L.
(Chicken).
tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
31;
[ N.A.
leghorn; TISSUE=Spleen;
Young J.R., Burnside J.;
putative chicken CD40 ligand";
P-2003) to the EMBL/GenBank/DBJ databases.
Cytokine that binds to INFRSFS. Mediates B-cell
ation in the absence of co-stimulus as well as IgE
n in the presence of IL-4. Involved in immunoglobulin
tching (By similarity).

```

```

CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists a
CC extracellular soluble form (By similarity).
CC -!- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a coll.
CC between the Swiss Institute of Bioinformatics and the EMBL out-
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for c-
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AJ243435; CAB95748.2; -.
CC HSP; P29965; ITALY.
CC GO; GO:0016021; C: integral to membrane; ISS.
CC GO; GO:0005174; F: CD40 receptor binding; ISS.
CC GO; GO:0042100; P: B-cell proliferation; ISS.
CC GO; GO:0006954; P: inflammatory response; ISS.
CC GO; GO:0007159; P: leukocyte cell adhesion; ISS.
CC GO; GO:0030168; P: platelet activation; ISS.
CC InterPro; IPR003263; TNF_5.
CC InterPro; IPR006052; TNF family.
CC InterPro; IPR008983; TNF-like.
CC InterPro; IPR003636; TNF_subf.
CC Pfam; PF00229; TNF; 1.
CC PRINTS; PR01702; CD40LIGAND.
CC ProDom; PD008600; TNF_5; 1.
CC ProDom; PD002012; TNF_subf; 1.
CC SMART; SM00207; TNF; 1.
CC PROSITE; PS00251; TNF_1; 1.
CC PROSITE; PS00049; TNF_2; 1.
CC Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
CC CHAIN 1 272
CC CHAIN 111 272
CC DOMAIN 1 23
CC TRANSMEM 24 44
CC DOMAIN 45 272
CC SITE 110 111
CC DISULFID 190 229
CC CARBOHYD 124 124
CC CARBOHYD 146 146
CC CARBOHYD 251 251
CC SEQUENCE 272 AA; 30832 MW; 8CD0338A924E044B CRC64;
Query Match 9.3%; Score 108.5; DB 1; Length 272;
Best Local Similarity 22.5%; Pred. No. 0.024;
Matches 58; Conservative 40; Mismatches 107; Indels 53; G
Qy 1 VLSGLALACGLLVVSLGSWATLSAQ-----FPSQELTAEDRRPE-----
Db 34 VQTIGTVFLCYLHKMKDKMEVLSLNEGYIFLRKVKQCTGDEQKSTLLDCEKYI
Qy 46 ELNPQTEESQDVVPFLQOLVPRPSAPKGRK-----ARPRRAIAAHYEVHPRPGOI
Db 94 DLQCKORTASEELPKFEMHGRHCHPHLKSNETSVAEKRPPIATHLA-----
Qy 101 GVDGTGVSQWEETK-INSSSPLRYDRQICEFTVIRAGLYLYCQVHFDEGKA-----
Db 146 NTVTRVLKMTTSYAPTSSLSIYHE--GKLKVEKAGLYIYSQVSFCTKAAASAPI
Qy 152 YLKLDDLVLNGVLAIRCLLEFSATASSPGPOLRCQV-----SGLLPLRPGSSLR.
Db 204 YLYLPMEDRL--MKGLDTHSTSTA-----LCELOSIREGGVVELRQGMDFV
Qy 207 WAHLKAAPFLTYFGLFQV 224
Db 255 STAVNVNPGNTYFGVFKL 272

```

STANDARD; PRT; 316 AA.
 6; Q9JJK8; Q9JJK9; Q9R1Y0;
 (Rel. 40, Created)
 (Rel. 40, Last sequence update)
 (Rel. 42, Last annotation update)
 is factor ligand superfamily member 11 (Receptor activator
 actor kappa B ligand) (RANKL) (TNF-related activation-
 ion factor) (TRANCE) (Osteoprotegerin ligand) (OPGL) (Osteoclast
 ion factor) (ODF) (Osteoclastogenesis-inhibitory factor)
 ANKL OR TRANCE OR OPGL.
 (Mouse).
 stazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 3090;
 M N.A. (ISOFORM 1).
 3112; PubMed=9312132;
 J., Arron J., Robinson E., Orlick J., Chao M.,
 ., Cayani E., Bartlett F.S. III, Frankel W.N., Lee S.Y.,
 novel ligand of the tumor necrosis factor receptor family
 as c-Jun N-terminal kinase in T cells.";
 n. 272:25190-25194(1997).
 M N.A. (ISOFORM 1).
 c lymphoma;
 2977; PubMed=9367155;
 ., Maraskovsky E., Billingsley W.L., Dougall W.C.,
 ., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,
 of the TNF receptor and its ligand enhance T-cell growth
 c-cell function.";
 75-179(1997).
 M N.A. (ISOFORM 1).
 marrow;
 7661; PubMed=9568710;
 Timms E., Tan H.-L., Kelley M.J., Dunstan C.R.,
 Elliott R., Colombero A., Elliott G., Scully S., Hsu H.,
 Hawkins N., Davy E., Capparelli C., Eli A., Qian Y.-X.,
 Sarosi I., Shalhoub V., Senaldi G., Guo J., Delaney J.,
 erin ligand is a cytokine that regulates osteoclast
 ion and activation.";
 176(1998).
 M N.A. (ISOFORM 1).
 marrow stroma;
 8248; PubMed=9520411;
 hima N., Nakagawa N., Yamaguchi K., Kinoshita M.,
 -I., Tomoyasu A., Yano K., Goto M., Murakami A., Tsuda E.,
 Higashio K., Udagawa N., Takahashi N., Suda T.;
 differentiation factor is a ligand for
 rin/osteoclastogenesis-inhibitory factor and is identical
 NK1.";
 Acad. Sci. U.S.A. 95:3597-3602(1998).
 M N.A. (ISOFORM 1).
 4075; PubMed=10196481;
 Kodaira K., Mizuno A., Yasuda H., Shima N., Murakami A.,
 (characterization of the gene encoding mouse osteoclast
 ion factor.";
 -127(1999).
 M N.A. (ISOFORMS 1; 2 AND 3).
 RX MEDLINE=21150053; PubMed=11250921;
 Ikeda T., Kasai M., Utsuyama M., Hirokawa K.;
 "Determination of three isoforms of the receptor activator of nu
 factor-kappaB ligand and their differential expression in bone a
 thymus.";
 Endocrinology 142:1419-1426 (2001).
 [7]
 RN SEQUENCE OF 139-147, PROCESSING, AND N-GLYCOSYLATION.
 RX MEDLINE=99240759; PubMed=10224132;
 Lum L., Wong B.R., Josien R., Becherer J.D., Erdjument-Bromage H
 RA Schleindorff J., Tempst P., Choi Y., Blobel C.P.;
 "Evidence for a role of a tumor necrosis factor-alpha
 (TNF-alpha)-converting enzyme-like protease in shedding of TRANC
 TNF family member involved in osteoclastogenesis and dendritic c
 survival.";
 J. Biol. Chem. 274:13613-13618(1999).
 RL [8]
 RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 158-316.
 RX MEDLINE=21464816; PubMed=11581298;
 Lam J., Nelson C.A., Ross F.P., Teitelbaum S.L., Fremont D.H.;
 "Crystal structure of the TRANCE/RANKL cytokine reveals determin
 of receptor-ligand specificity.";
 J. Clin. Invest. 108:971-979(2001).
 RL [9]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 137-316.
 RX MEDLINE=21839021; PubMed=11733492;
 Ito S., Wakabayashi K., Ubukata O., Hayashi S., Okada F., Hata T
 RA "Crystal structure of the extracellular domain of mouse RANK lig
 RT 2.2-A resolution.";
 J. Biol. Chem. 277:6631-6636(2002).
 RL CC -I- FUNCTION: Cytokine that binds to TNFRSF11B/OPG and to
 CC TNFRSF11A/RANK. Osteoclast differentiation and activation fa
 CC Augments the ability of dendritic cells to stimulate naive T
 CC proliferation. May be an important regulator of interactions
 CC between T cells and dendritic cells and may play a role in t
 CC regulation of the T cell-dependent immune response. May also
 CC an important role in enhanced bone-resorption in humoral
 CC hypercalcemia of malignancy.
 CC -I- SUBUNIT: Homotrimer.
 CC -I- SUBCELLULAR LOCATION: Type II membrane protein and secreted
 CC (isoforms 1 and 2); Cytoplasmic (isoform 3).
 CC -I- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=O35235-1; Sequence=VSP_006449;
 CC Name=2;
 CC IsoId=O35235-2; Sequence=VSP_006449;
 CC Name=3;
 CC IsoId=O35235-3; Sequence=VSP_006448;
 CC -I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THYMUS AND LYMPH NOD
 CC NOT IN NONLYMPHOID TISSUES AND IS ABUNDANTLY EXPRESSED IN T
 CC BUT NOT IN B CELLS. A HIGH LEVEL EXPRESSION IS ALSO SEEN IN
 CC TRABECULAR BONE AND LUNG.
 CC -I- PTM: N-glycosylated.
 CC -I- PTM: The soluble form of isoform 1 derives from the membrane
 CC by proteolytic processing. The cleavage may be catalyzed by
 CC ADAM17. A further shorter soluble form was observed.
 CC -I- DISEASE: Deficiency in TNFSF11 results in failure to form lc
 CC alveolar mammary structures during pregnancy, resulting in c
 CC of newborns. Trance-deficient mice show severe osteopetrosis
 CC no osteoclasts, marrow spaces, or tooth eruption, and exhibi
 CC profound growth retardation at several skeletal sites, inclu
 CC the limbs, skull, and vertebrae and have marked chondrodyspl
 CC with thick, irregular growth plates and a relative increase
 CC hypertrophic chondrocytes.
 CC -I- SIMILARITY: Belongs to the tumor necrosis factor family.
 CC This SWISS-PROT entry is copyright. It is produced through a col
 CC between the Swiss Institute of Bioinformatics and the EMBL o
 CC the European Bioinformatics Institute. There are no restricti
 CC use by non-profit institutions as long as its content is i
 CC modified and this statement is not removed. Usage by and for
 CC entities requires a license agreement (See http://www.isb-sib.ch

```

il to license@isb-sib.ch).
; AAC71061.1; -
; AAB86812.1; -
; AAC40113.1; -
; BAA25425.1; -
; BAA36970.1; -
; BAA36970.1; JOINED.
; BAA36970.1; JOINED.
; BAA36970.1; JOINED.
; BAA97257.1; -
; BAA97258.1; -
; BAA97259.1; -
JAN-03.
JAN-03.
; 89; tnfs11.
; P:organogenesis; IMP.
; P:ossification; IMP.
006052; TNF family.
008983; TNF-like.
003636; TNF_subf.
; TNF; 1.
012; TNF_subf; 1.
7; TNF; 1.
251; TNF; 1; FALSE_NEG.
049; TNF_2; 1.
fermentation; Receptor; Glycoprotein; Transmembrane;
; 3D-structure; Alternative splicing
1 316
39 316
1 48
49 69
CYTOPLASMIC (POTENTIAL).
SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
EXTRACELLULAR (POTENTIAL).
CLEAVAGE.
38 139
97 197
62 262
N-LINKED (GLCNAC. .) (POTENTIAL).
Missing (in isoform 3).
1 117
/FTID=VSP 006448.
SSEENSGSGVPGHPEGLHPAPSAPAPPPA -> TP (in
isoform 2).
/FTID=VSP 006449.
G -> D (IN REF. 2).
MISSING (IN REF. 5).

99 99
41 143
64 169
71 172
81 182
86 187
91 192
94 196
201 98
202 203
204 207
211 224
225 227
234 245

8.8; Score 102.5; DB 1; Length 316;
arity 24.3%; Pred. No. 0.096;
Conservative 37; Mismatches 107; Indels 65; Gaps 13;

ALACGLLVVYSLGSAWTLAQ-BPSQELTAEDR-----RPPPELNPT 51
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
ALLGLGLGVVCSIALFLYFRAQMDPNR---ISEDTHCFYRLRLHFNAGLQDST 104
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QDVVP-----FLEQLVRRPR--SAPK-----GRKARPRRAA 84
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EDTLFDSCRMKQAFQAVQKELQHVGPQFSGAPAMWGSWLDVAQRGKPEAQPF 164
EVHPRPQDGAQGVGDVTSGWEE-----TKINSSPLRYDRQIGFTVIRAGLYLY 140
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5

```

TNFS_CANFA STANDARD; PRT; 260 AA.
ID TNFS_CANFA
AC 097626;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand).
GN TNFSF5 OR CD40LG OR CD40L.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Hosie M.H., Willett B.J.;
RT "Adjuvant properties of canine CD40L.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
CC proliferation in the absence of IL-4. Involved in immunoglobulin
CC production in the presence of IL-4.
CC Class switching (By similarity).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as
CC extracellular soluble form (By similarity).
CC -!- PM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
-----
This SWISS-PROT entry is copyright. It is produced through a coll
between the Swiss Institute of Bioinformatics and the EMBL out
the European Bioinformatics Institute. There are no restriction
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for
entities requires a license agreement (See http://www.isb-sib.ch,
or send an email to license@isb-sib.ch).
-----
EMBL; AF086711; AAD04375.1; -.
HSSP; P29965; ITALY.
GO; GO:0016021; C: integral to membrane; ISS.
GO; GO:0005174; F: CD40 receptor binding; ISS.
GO; GO:0042100; P: B-cell proliferation; ISS.
GO; GO:0006954; P: inflammatory response; ISS.
GO; GO:0007159; P: leukocyte cell adhesion; ISS.
GO; GO:0030168; P: platelet activation; ISS.
InterPro; IPR003263; TNF_5.
InterPro; IPR006052; TNF family.
InterPro; IPR008983; TNF like.
InterPro; IPR003636; TNF_subf.
Pfam; PF00229; TNF; 1.
PRINTS; PR01702; CD40LIGAND.
ProDom; PD008600; TNF_5; 1.
ProDom; PD002012; TNF_subf; 1.
SMART; SM00207; TNF; 1.
PROSITE; PS00251; TNF_1; 1.
PROSITE; PS50049; TNF_2; 1.
Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
TUMOR NECROSIS FACTOR LIGAND SUPERF
MEMBER 5, MEMBRANE FORM.
TUMOR NECROSIS FACTOR LIGAND SUPERF
MEMBER 5, SOLUBLE FORM (BY SIMILARI
CYTOPLASMIC (POTENTIAL).
CHAIN 1 260
CHAIN 112 260
DOMAIN 1 22

```

	proteolytic processing (By similarity)).
-!	SIMILARITY: Belongs to the tumor necrosis factor family.
This	SWISS-PROT entry is copyright. It is produced through a coll-
between	the Swiss Institute of Bioinformatics and the EMBL out-
the	European Bioinformatics Institute. There are no restriction
use	by non-profit institutions as long as its content is ir-
modified	and this statement is not removed. Usage by and for c-
entities	requires a license agreement (See http://www.ebi-sib.ch/
or send	an email to licensese@ebi-sib.ch).
EMBL; AF187319;	AAG17031.1; -
EMBL; AF425669;	AAL23963.1; -
HSSP; P50591;	IDOG.
InterPro; IPR006052;	TNF family.
InterPro; IPR008983;	TNF like.
InterPro; IPR003636;	TNF_subf.
Pfam; PF00229;	TNF_1
ProDom; PD002012;	TNF_subf; 1.
SMART; SMO0207;	TNF; 1.
PROSITE; PS00251;	TNF_1; FALSE_NEG.
PROSITE; PS00049;	TNF_2; 1.
Cytokine; Differentiation;	Receptor; Glycoprotein; Transmembrane
Signal-anchor.	1 318 TUMOR NECROSIS FACTOR LIGAND SUPERF;
CHAIN	MEMBER 11, MEMBRANE FORM.
CHAIN	141 318 TUMOR NECROSIS FACTOR LIGAND SUPERF;
DOMAIN	1 47 CYTOPLASMIC (POTENTIAL).
TRANSMEM	48 68 SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO
FT FT	(POTENTIAL).
DOMAIN	69 318 EXTRACELLULAR (POTENTIAL).
SITE	140 141 CLEAVAGE (BY SIMILARITY).
CARBOHYD	199 199 N-LINKED (GLCNAC..)
CARBOHYD	264 264 N-LINKED (GLCNAC..)
CONFLICT	317 317 I->M (IN REF. 2).
SEQUENCE	318 AA; 35370 MW; 487A4D706AD098F CRC64;
Query Match	7.8%; Score 90.5; DB 1; Length 318;
Best Local Similarity	23.3%; Pred.No.1.1;
Matches	64; Conservative 36; Mismatches 108; Indels 67;
QY	6 LALACLGLLLVVSLGSAWATLSAQ-PSPQEELTAEDR-----REPPELNPI
Db	51 LALLGLGLGVCCSIALLYFLRYQMPPNR---ISEDSTRCFYRIILEKRENTGLQD.
QY	55 QD-----VVFPLEQLVRPRESA.....PKGRKAHPE
Db	108 ETETALPDSCRMKQAFCQGAVORELIHVGPQRFSVPAMMGSLWDVARRGKPE
QY	86 HYEHPRPFGODGAQGDVTGSWEIE-----TKINSSTPLYRDQIGEFTVIIRAGL
Db	168 HLTIINAIDPSGSH--KVSLSNWSYHDRGWAKISNMT-----LSNGKLRFVNQDFG
QY	142 QV----HFDEKK-AAYIKLDLVV---NGVLALRCLEFSATSATASSPGPOLRL-
Db	220 NICFRHHETSIVPADYLQMLVYYVKTSIKIPSHNLMKGGSTKNWGNSGFPHY
QY	190 GLLPRLPCSGSRIITLPAWHLKAAFFLTYPGLFOV 224
Db	280 GFKKURAGEEISVSQNPSLLDDPDQDYTGAFKV 314
RESULT 7	
PPEN_HUMAN	
ID_PPEN_HUMAN	STANDARD; PRT; 532 AA.
AC	F106596; Q16727; Q96CWI;
DT	01-APR-1990 (Rel. 14, Created)
DT	01-JUN-1994 (Rel. 23, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Alkaline phosphatase, placental-like precursor (EC 3.1.3.1) (Naq
GN	isozyme) (German-cell alkaline phosphatase) (GCAP) (PLAP-like) (AL
DN	APP2 OR ALPLP.

Human).
Chordata; Vertebrata; Euteleostomi;
Pisces; Catarrhini; Hominidae; Homo.
5;

N.A.
79; PubMed=2162249;
auss A.W.;
a Nagao-type, phosphatidylinositol-glycan anchored
thase in human choriocarcinomas.";
:3956-3962(1990).

N.A.
 carcinoma;
 96; PubMed=2745460;
 atanabe T., Li W.L., Soong B.-W., Chou J.Y.;
 the germ cell alkaline phosphatase gene in human
 a cells.";
 264:12611-12619(1989).

N.A.
32; PubMed=2834730;
anes T.;
ved Nagao isozyme is encoded by a germ-cell alkaline
ne."
ad Sci. U.S.A. 85:3024-3028(1988).

N.A.
11; PubMed=2297757;
S J.W., Sack T.L., Kim Y.S.;
ning of complementary DNAs encoding alkaline
human colon cancer cells.";
:1085-1091(1990).

N.A.
a;
57; PubMed=12477932;
Collings E.A., Grouse L.H., Derge J.G.,
Feinberg F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Zeeberg B., Moore K.H., Schaefer C.F., Bhat N.K.,
Jordan H., Bute T., Max S.I., Wang J., Hsieh F.,
Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.,
Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Ruellan N.A., Peters G.J., Abramson R.D., Mullany S.,
Ewan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Forley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.,
Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Kettner M., Madaen A., Rodrigues S., Sanchez
dan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Touchman J.W., Green E.D., Dickson M.C.,
Grinnwood J., Schmutz J., Myers R.M.,
S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
Schein J.E., Jones S.J.M., Marra M.A.;
initial analysis of more than 15,000 full-length
cDNA sequences.”
ad. Sci. U.S.A. 99:16899-16903 (2002).

157 FROM N.A.
378; PubMed=3387245;
1 H., Kan Y.W., Kam W.;
the sequence of a putative human placental alkaline
like gene.";

ACTIVITY: An orthophosphoric monoester + H₂O = an
phosphate.
Monomer.

-- SIMILARITY: Belongs to the alkaline phosphatase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outposts at the European Bioinformatics Institute. There are no restrictions on the use of this information as long as its content is in the public domain. This statement is not removed, Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/2>) or send an email to license@isb-sib.ch).

EMBL	X55958	CAA39425.1	-
EMBL	J04948	AAA51700.1	-
EMBL	J03252	AAA98616.1	-
EMBL	X53279	CAA37374.1	-

EMBL; SC041139; GENE333.1; ALT_SEQ.
DR X07247; CAA30232.1; ALT_SEQ.
PIR; S12076; S12076.
DR

DR Sienra-2DPAGE; P10696; -
DR Genew; HGNC:441; ALPPL2.

DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0004035; F:alkaline phosphatase activity; NAS.

DR InterPro; IPR001952; Alk_phosphatse.
DR Pfam; PF00245; alk_phosphatase; 1.
DR PROSITE; PS00112; ALKPHOSPHATASE.

DR ProDom; PD001868; Alk_phosphatase; 1.
DR SMART; SM00098; alkPPC; 1.
DR DPGSITE; DS00123; ALKALINE PHOSPHATASE. 1

KW Hydrolase; Zinc; Magnesium; Phosphorylation; Transmembrane;
KW Multigene family; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL. 1 19 POTENTIAL.

FT	CHAIN	20	303	ALKALINE PHOSPHATASE, PLACENTAL
FT	PROPEP	504	532	REMOVED IN MATURE FORM.
FT	DISULFID	140	202	BY SIMILARITY.

FT	100	DISOLVD	100	2.5	PHOSPHORINE INTERMEDIATE.
FT	111	ACT SITE	111	111	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	141	CARBOHYD	141	141	

FT	57	I -> M (IN REF. 3).
CONFLICT	152	V -> M (IN REF. 2, 4, 5 AND 6).

PT	CONFLICT	260	H -> R (IN REF. 5).
PT	CONFLICT	273	L -> M (IN REF. 4 AND 5).

FT	CONFLICT	380	V → L (IN REF. 2)	380
FT	CONFLICT	498	P → R (IN REF. 2 AND 5)	498
FT	CONFLICT	498	D → S (IN REF. 4)	498

FT	531	CONFLICT	531	A -> T (IN REF. 4).
SQ	532	SEQUENCE	532	AA; 57315 MW; 84AB5B28F13D6D82 CRC64;

Query Matchn 7.6%; Score 88; DS 1; Length 332;
Best Local Similarity 28.6%; Pred. No. 3.4;
Matches 44; Conservative 16; Mismatches 44; Indels 50; G:

QY 76 KAPRRRAIAHYEV---HPRPG---QDGAQAGVDGTVSGWEETKINSSPLRYDRQ.

QY 130 TVIRAGLYILYCQVHFDEGKAVYIKDLLVNGV-----LALRCLEEFSA--

Db 456 VAV-----LVHGVQETFLAHVMAFAACLEPYTACD;

Db 499 AGTTDAAHPGGSV----VPALLPLLAGTLLLLGT 528

RESULT 8

ID MINT_HUMAN STANDARD; PRT; 3664 AA.


```

0 325 SER-RICH.
6 810 ARG-RICH.
4 697 TYR-RICH.
8 2520 PRO-RICH.
0 3482 PRO-RICH.
0 970 A -> V (in dbSNP:848208).
1 1091 /FTId=VAR_017119.
0 2360 L -> P (in dbSNP:848209).
0 956 /FTId=VAR_017120.
0 2360 N -> D (in dbSNP:848210).
6 956 /FTId=VAR_017121.
4 AA; 402245 MW; 52280585335B27B CRC64;
7.6%; Score 88; DB 1; Length 3664;
urity 23.1%; Pred.No. 35;
nservative 20; Mismatches 62; Indels 58; Gaps 7;
'AADREPPPELNPQTEESQ-----DVVPFLEQLVRRRSAPK----- 73
'AGPANRPEPHTQVQRAAETGTFSPSPVSMKPD-L-PVSLPTQTAPKQPLFV 3459
'KARPRATAAHVEVHPRGQDCAQAGVDGTVSGWEETKINSPLRYDRQIGFT 130
'PSTPEGLVLPHTFQFPAPKQDSS-----PHLTSQRPVDMVQLLKXP 3506
HLYL-----YCOVHDEGKAVYLKDLLVNGVIALRCLLEFSATAASPGPOLRL 185
HLLALNDTAAVQLHFVSG-----NNVLAHRSL-----PLSEGGPLRI 3549
37
551

; STANDARD; PRT; 244 AA.
); Q99761;
rel. 29, Created)
rel. 29, Last sequence update)
rel. 43, Last annotation update)
ta (IT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
OR TNFC.
(Human).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
);
N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
381; PubMed=7916655;
Ngam-Ek A., Lawton P., Demarinis J., Tizard R.,
sson C., O'Brine-Greco B., Foley S.F., Ware C.F.;
beta, a novel member of the TNF family that forms a
complex with lymphotoxin on the cell surface.";
56(1993).
N.A. (ISOFORMS 1 AND 2).
965; PubMed=9299492;
Renard N., Charlot C., Bienvenu J., Coiffier B.,
on of two lymphotoxin beta isoforms expressed in human
lines and non-Hodgkin's lymphomas.";
lys. Res. Commun. 238:273-276 (1997).
N.A. (ISOFORM 1).
Malner C.M., Campbell R.D.;
of the immunoglobulin superfamily and a V-ATPase G
mongst the predicted products of novel genes close to the
the human MHC.";
P-1997) to the EMBL/GenBank/DBJ databases.

```

```

RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
RA Lasky S., Hood L.;
RT "Sequence of the human major histocompatibility complex class III
RT region.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Shiina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLU-70 AND
RP PRO-111.
RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS ARG-84 AND PHE-87.
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytokine that binds to LTBR/TNFRSF3. May play a spe
CC role in immune response regulation. Provides the membrane anc
CC for the attachment of the heterotrimeric complex to the cell
CC surface. Isoform 2 is probably non-functional.
CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
CC (less prevalent) one LTB and two LTA subunits.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q06643-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q06643-2; Sequence=VSP_006441, VSP_006442;
CC -1- TISSUE SPECIFICITY: Spleen and thymus.
CC -1- SIMILARITY: Belongs to the tumor necrosis factor family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a coll
CC between the Swiss Institute of Bioinformatics and the EMBL out
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for c
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L11016; AAA99888.1; -
EMBL; U89922; AAC51769.1; -
EMBL; U79029; AAB37342.1; -
EMBL; L11015; AAA36191.1; -
EMBL; Y14768; CAA75069.1; -
EMBL; AF129756; AAD18089.1; -
EMBL; AP000505; BAB63395.1; -
EMBL; AY070219; AAL49954.1; -
EMBL; AY070219; AAL49955.1; -
EMBL; AY216497; AAO21134.1; -
EMBL; A46066; A46066.
EMBL; JCS645; JCS645.
EMBL; P01374; ITNR.
EMBL; HGNC:6711; LTB.
EMBL; MIM: 600978; -.
EMBL; GO:0005102; F:receptor binding; TAS.
EMBL; GO:0015070; P:toxin activity; NAS.
EMBL; GO:0007267; P:cell-cell signaling; TAS.
EMBL; GO:0007165; P:signal transduction; TAS.
EMBL; InterPro; IPR006053; TNF abc.
EMBL; InterPro; IPR006052; TNF family.
EMBL; InterPro; IPR008983; TNF-like.
EMBL; InterPro; IPR003636; TNF_subf.
EMBL; Pfam; PF00229; TNF; 1.

```

```

34; TNECROSISFCT.
1012; TNF subf; 1.
17; TNF; 1.
1251; TNF_1; 1.
1049; TNF_2; 1.
membrane; Glycoprotein; Signal-anchor;
splicing; Polymorphism.
1 18 CYTOPLASMIC (POTENTIAL).
19 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
49 244 EXTRACELLULAR (POTENTIAL).
122 222 N-LINKED (GLCNAC... ) (POTENTIAL).
53 77 GLVTETADPGAQOGLGFKLPEE -> GLGFRSCQRRSQ
KQISAPGSQLPTS (in isoform 2).
/FTId-VSP 006441.
Missing (in isoform 2).
78 244 /FTId-VSP 006442.
70 70 G->E_013025.
84 84 S->R_013025.
87 87 /FTId-VAR_016331.
L->F_016332.
11 111 /FTId-VAR_016332.
A>P_013026.
60 69 /FTId-VAR_013026.
DPAQAQOGL -> GLSAPGSGRT (IN REF. 2;
AAB37342).
14 AA; 25390 MW; F41569459830ED4C CRC64;
7.5%; Score 87; DB 1; Length 244;
arity 23.0%; Pred. No. 1.6;
Conservative 26; Mismatches 80; Indels 108; Gaps 13;
ALA---CLGLLVVSLGSWATLSAQEPSQBELTAEDR-----REPPPEL 47
AVAGATSLVTLIAVPIITVLAVLVPDQOGLVTETADPGAQAQOGLGFKLPEE 77
TESQDVVPLEQLVPRRSAPKGRKARRAIAAHYVHPREGQDGAQAGVDGTVS 107
T-----DLSPGLP-----AAHLIGAPLKGQ-----L 102
ETKINS--SSPLRYDRQIGFTVIRAGLYLYCQVHF-----DEKAVYLKL 155
TKEQAFLTSGTQFSDAEG-LALPDQGLLYLYCLVGRAPPGGGDPQGRSVTLRS 161
NGVLALRCLEEFSAATAASPG-PQLRL-----CQVS 189
-----YRAGGAYGPGTPELLLEGAETVTPVLPDARRQGYGLMYTSVGFG 208
PLRPGSLRIRTLPAHLKAAPFL---TYFGLFQV 224
LRRGERVYVNI---SHPDMVDFAFGKTFPGAVMV 243
STANDARD; PRT; 244 AA.
(Rel. 43, Created)
(Rel. 43, Last sequence update)
(Rel. 43, Last annotation update)
beta (IIT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
ligand superfamily member 3).
3 OR TNFC.
tes (Chimpanzee).
atazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Primates; Catarrhini; Homnidae; Pan.
598;
M N.A.
1002; PubMed=12493009;
Shiina T., Anzai T., Kohara S., Inoko H.;
genomic analysis of the MHC: the evolution of class I
RT duplication blocks, diversity and complexity from shark to man."
RN Immunol. Rev. 190:95-122(2002).
RP SEQUENCE FROM N.A.
RX MEDLINE=22709134; PubMed=12799463;
RA Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenar:
RA Yamagata T., Kuleki J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
RA Yamazaki M., Tashiro H., Iwamoto C., Umehara Y., Imanishi T.,
RA Meyer A., Ikeo K., Gojobori T., Bahram S., Inoko H.;
RT "Comparative sequencing of human and chimpanzee MHC class I regi:
RT unveils insertions/deletions as the major path to genomic
RT divergence."
RT Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713(2003).
RL -!- FUNCTION: Cytokine that binds to LTBR/TNFRSF3. May play a sp:
CC role in immune response regulation. Provides the membrane an:
CC for the attachment of the heterotrimeric complex to the cell
CC surface (By similarity).
CC -!- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
CC (less prevalent) two LTA and one LTB subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a col:
CC between the Swiss Institute of Bioinformatics and the EMBL ou:
CC the European Bioinformatics Institute. There are no restriction:
CC use by non-profit institutions as long as its content is u:
CC modified and this statement is not removed. Usage by and for:
CC entities requires a license agreement (See http://www.isb-sib.ch
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB054536; BAB3881.1; -;
CC EMBL; AB100082; BAC78156.1; -;
CC InterPro; IPR006053; TNF abc.
CC InterPro; IPR006052; TNF family.
CC InterPro; IPR008983; TNF like.
CC InterPro; IPR003636; TNF subf.
CC Pfam; PF00229; TNF; 1.
CC PRINTS; PR01234; TNECROSISFCT.
CC ProDom; PD002012; TNF subf; 1.
CC SMART; SM00207; TNF; 1.
CC PROSITE; PS00251; TNF_1; 1.
CC PROSITE; PS50049; TNF_2; 1.
CC Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
KW CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 18 SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO
FT TRANSMEM 19 48 (POTENTIAL).
FT DOMAIN 49 244 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 222 222 N-LINKED (GLCNAC... ) (POTENTIAL).
FT SEQUENCE 244 AA; 25420 MW; A4047858335DSB97 CRC64;
SQ
Query Match 7.5%; Score 87; DB 1; Length 244;
Best Local Similarity 23.0%; Pred. No. 1.6;
Matches 64; Conservative 26; Mismatches 80; Indels 108;
QY 3 SLGLALA---CLGLLVVSLGSWATLSAQEPSQBELTAEDR-----R
DB 18 SLLAVAGATSLVTLIAVPIITVLAVLVPDQOGLVTETADPGAQAQOGLGFG
QY 48 NPQTEESQDVVPLEQLVPRRSAPKGRKARRAIAAHYVHPREGQDGAQAGV
DB 78 EPEE---DLSPGLP-----AAHLIGAPLKGQ-----
QY 108 GWETTKINS--SSPLRYDRQIGFTVIRAGLYLYCQVHF-----DSGKA
DB 103 GWETTKQAFLTSGTQFSDAEG-LALPDQGLLYLYCLVGRGRTPTPPGGDPQGRS
QY 156 DLLVNGVLALRCLEEFSAATAASPG-PQLRL-----
DB 162 SL-----YRAGGAYGPGTPELLLEGAETVTPVLPDARRQGYGLMYT
QY 190 GLLPRLRPGSLRIRTLPAHLKAAPFL---TYFGLFQV 224
DB 209 GLVQLRRGERVYVNI---SHPDMVDFAFGKTFPGAVMV 243

```

STANDARD; PRT; 261 AA.

el. 41, Created)
el. 41, Last sequence update)
el. 41, Last annotation update)
factor ligand superfamily member 5 (CD40 ligand) (CD40-
ein).
LG OR CD40L.
tus (Night monkey) (Douroucouli).
aza; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
5;

N.A.

ytes;

18; PubMed=11491535;

Bostik P., Mayne A.E., King C.L., Genain C.P.,

sari A.A.;

encing, and homology analysis of nonhuman primate
and co-stimulatory molecules.";

53:315-328(2001).

Cytokine that binds to TNFRSF5. Mediates B-cell
ion in the absence of co-stimulus as well as IGE
in the presence of IL-4. Involved in immunoglobulin
ching (By similarity).

omotrimer (By similarity).

R LOCATION: Type II membrane protein. Also exists as an
lar soluble form (By similarity).

oluble form derives from the membrane form by

c processing (By similarity).

: Belongs to the tumor necrosis factor family.

T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See <http://www.isb-sib.ch/announce/>
il to license@isb-sib.ch).

AAK37542.1;

1ALY.

; C-integral to membrane; ISS.

; F:CD40 receptor binding; ISS.

; P:B-cell proliferation; ISS.

; P:inflammatory response; ISS.

; P:leukocyte cell adhesion; ISS.

; P:platelet activation; ISS.

03263; TNF_5.

06052; TNF_family.

08983; TNF_like.

03636; TNF_subf.

TNF; 1.

2; CD40LIGAND.

00; TNF_5; 1.

12; TNF_subf; 1.

; TNF; 1.

51; TNF_1; 1.

49; TNF_2; 1.

membrane; Glycoprotein; Signal-anchor.

1 261

TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

MEMBER 5, MEMBRANE FORM.

3 261

TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

MEMBER 5, SOLUBLE FORM (BY SIMILARITY).

1 22

CYTOPLASMIC (POTENTIAL).

3 43

SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)

(POTENTIAL).

4 261

EXTRACELLULAR (POTENTIAL).

2 113

CLEAVAGE (BY SIMILARITY).

FT DISULFID 178 218 POTENTIAL.
FT CARBOHYD 240 240 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 261 AA; 29357 MW; 85E1588B507901B5 CRC64;

Query Match 7.5%; Score 87; DB 1; Length 261;
Best Local Similarity 25.9%; Pred. No. 1.8;
Matches 35; Conservative 21; Mismatches 55; Indels 24; Gs

QY 105 TVSGWEE---TKINSSSPLRYDQIGFTVIRAGLYLYCOVHFDEGKAVYLKLDI
Db 136 SVLQWAEKGYTMSNNVLTLENGKQL---TVKRGLYIYIAQVTCNSNEASSQAPF

QY 161 GVALRLCLEEF-----SATAASSPQQLRLC-----QVSGLLPLRPSSSLRIRTLF
Db 193 --LCLKPPNRFERILLRAANTHSSAKP-----CGQOSIHLLGGIFELQPGASVFNVTI

QY 210 LKAAPFLTYFGLFQV 224

Db 247 VSHGTGFTSFGLLKL 261

RESULT 12

TNFS_CALJA

ID TNFS_CALJA STANDARD; PRT; 261 AA.

AC Q9BDN3;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (C
L) (CD154 protein).

GN TNFSF5 OR CD40LG OR CD40L.

OS Callithrix jacchus (Common marmoset).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;

OC Callithrix.

OX NCBI_TaxID=9483;

RN [1]

SEQUENCE FROM N.A.

RC TISSUE=Lymphocytes;

RX MEDLINE=21383618; PubMed=11491535;

RA Willinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,

Weiss W.R., Ansari A.A.;

"Cloning, sequencing, and homology analysis of nonhuman primate

Fas/Fas-ligand and co-stimulatory molecules.";

RL Immunogenetics 53:315-328(2001).

CC -!- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
proliferation in the absence of co-stimulus as well as IGE

production in the presence of IL-4. Involved in immunoglobulin
class switching (By similarity).

CC -!- SUBUNIT: Homotrimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as
extracellular soluble form (By similarity).

CC -!- PTM: The soluble form derives from the membrane form by

proteolytic processing (By similarity).

CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.

CC This SWISS-PROT entry is copyright. It is produced through a colla
between the Swiss Institute of Bioinformatics and the EMBL outst
the European Bioinformatics Institute. There are no restrictions
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for co
entities requires a license agreement (See <http://www.isb-sib.ch/>
or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF344844; AAK37603.1; --

DR HSSP; P29965; 1ALY.

DR GO; GO:0016021; C:integral to membrane; ISS.

DR GO; GO:0005174; F:CD40 receptor binding; ISS.

DR GO; GO:0042100; P:B-cell proliferation; ISS.

DR GO; GO:0006954; P:inflammatory response; ISS.

DR GO; GO:0007159; P:leukocyte cell adhesion; ISS.

DR GO; GO:0030166; P:platelet activation; ISS.

DR InterPro; IPR003263; TNF_5.

30 950 POTENTIAL.
 34 1006 POTENTIAL.
 37 1053 POTENTIAL.
 37 1132 POTENTIAL.
 59 1195 POTENTIAL.
 59 1195 PRO-RICH.
 73 87 HIS-RICH.
 73 87 POLY-SER.
 61 865 N-LINKED (GLNAC. . .) (POTENTIAL).
 55 855 N-LINKED (GLNAC. . .) (POTENTIAL).
 66 866 N-LINKED (GLNAC. . .) (POTENTIAL).
 78 878 N-LINKED (GLNAC. . .) (POTENTIAL).
 69 1169 S-palmitoyl cysteine (By similarity).
 1 17 MSSAPRRPASGADSLHT -> MDLLRPQ (in isoform B2).
 1 17 /FTId=VSP 000457.
 1 17 MSSAPRRPASGADSLHT -> MTQ (in isoform B1).
 1 17 /FTId=VSP 000458.
 1 166 Missing (in isoform C2).
 1 198 /FTId=VSP 000459.
 1 198 Missing (in isoform C1).
 1 198 /FTId=VSP 000460.
 67 193 ERTSPPTPTPHQAPRAASKAQTG -> MPAPQEWKSG
 1 193 GLREAVFGAGCSVCR (in isoform C2).
 1 205 /FTId=VSP 000461.
 1 205 A -> G (in REF. 2).
 37 AA; 136813 MW; 1A0782C0071782EE CRC64;
 7 4%; Score 86.5; DB 1; Length 1237;
 arity 27.9%; Pred. No. 13;
 conservative 19; Mismatches 77; Indels 51; Gaps 12;
 TLQAQPSQBELT-AEDRPEPELNPQTESQDVVPLEQ-----LVRPRRSAPKG 74
 TIERGEDEERASEAGFRAPQ-QPSATTPSAVQFLQDEGAERKPERTSPSP 173
 PRRAIAHYVHPRGQCAQGV-----DGTVSGHEETKI---N 115
 "P-----HOEAAPRASK-GAQTGTIVEEMVAVASATAGDGGAGRLTKAQPCH 226
 "LYDROICEFTVIRAGLYLVCOVHFDGKAVYL---KLDLL-----VNGVL-- 163
 "LQERRIGSMGTVEQA---LPRVPTDSEAQTLATADLDLMLKSHRFDVGVRRH 283
 "LEFSATAA---SSPGQLR 184
 "NAKGSTQAAREGREGPTPR 307
 STANDARD; PRT; 261 AA.
 (Rel. 25, Created)
 (Rel. 25, Last sequence update)
 (Rel. 43, Last annotation update)
 s factor ligand superfamily member 5 (CD40 ligand) (CD40-
 ed activation protein) (TRAP) (T cell antigen Gp39)
 an).
 (OLG OR CD40L OR TRAP.
 (Human).
 atazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 theria; Primates; Catarrhini; Hominidae; Homo.
 506;
 N A.
 6854; PubMed=1280226;
 chauer U., Mages H.W., Senger G., Kroczeck R.A.;
 TRAP, a ligand for CD40 on human T cells.";
 nol. 22:3191-3194(1992).
 N A.
 9181; PubMed=1385114;
 D., Grosmaire L.S., Kullas C.D., Chalupny J.N.,

RA Braesch-Andersen S., Noelle R.J., Stamenkovic I., Ledbetter J.A.,
 RA Aruffo A.;
 RA "The human T cell antigen gp39, a member of the TNF gene family,
 RT ligand for the CD40 receptor: expression of a soluble form of gp3
 with B cell co-stimulatory activity.";
 RL EMBO J. 11:4313-4321(1992).
 [3]
 RN SEQUENCE FROM N.A., AND VARIANTS HIGM1 128-ARG-GLY-129 AND PRO-23
 RX MEDLINE=931145330; PubMed=7678782;
 RA Aruffo A., Farrington M., Hollenbaugh D., Li X., Milatovich A.,
 RA Novoyama S., Bajorath J., Grosmaire L.S., Stenkamp R., Neubauer M
 RA Roberts R.L., Noelle R.J., Ledbetter J.A., Francke U., Ochs H.D.;
 RA "The CD40 ligand, gp39, is defective in activated T cells from
 RT patients with X-linked hyper-IgM syndrome.";
 RL Cell 72:291-300(1993).
 [4]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93094757; PubMed=1281209;
 RA Spriggs M.K., Amitage R.J., Strockbine L., Clifford K.N.,
 RA Macduff B.M., Sato T.A., Maliszewski C.R., Fanslow W.C.;
 RA "Recombinant human CD40 ligand stimulates B cell proliferation an
 RT immunoglobulin E secretion.";
 RL J. Exp. Med. 176:1543-1550(1992).
 [5]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93138085; PubMed=7678552;
 RA Gauchat J.F.M., Aubry J.-P., Mazzei G.J., Life P., Jomotte T.,
 RA Elson G., Bonnefoy J.Y.;
 RT "Human CD40-ligand: molecular cloning, cellular distribution and
 RT regulation of expression by factors controlling IgE production.";
 RL FEBS Lett. 315:259-266(1993).
 [6]
 RN SEQUENCE FROM N.A.
 RA Shimadzu M., Terasaki H., Ninomiya R., Shimizu S., Nunoi H.,
 RA Matsuoka I.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 [7]
 RN SEQUENCE OF 113-117, AND PROCESSING.
 RX MEDLINE=96198042; PubMed=8626375;
 RA Pietravalle F., Leconet-Henchoz S., Blasey H., Aubry J.-P., Els
 RA Edgerton M.D., Bonnefoy J.-Y., Gauchat J.-P.;
 RT "Human native soluble CD40L is a biologically active trimer, pro
 RT inside microsomes.";
 RL J. Biol. Chem. 271:5965-5967(1996).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 116-261.
 RX MEDLINE=96131874; PubMed=8589998;
 RA Karpusas M., Hsu Y.-M., Wang J.-H., Thompson J., Lederman S.,
 RA Chess L., Thomas D.;
 RT "2-A crystal structure of an extracellular fragment of human CD4
 RT ligand.";
 RL Structure 3:1031-1039(1995).
 [9]
 RN 3D-STRUCTURE MODELING OF COMPLEX WITH CD40.
 RX MEDLINE=98266353; PubMed=9605317;
 RA Singh J., Garber E., van Vlijmen H., Karpusas M., Hsu Y.-M.,
 RA Zheng Z., Naismith J.H., Thomas D.;
 RT "The role of polar interactions in the molecular recognition of
 RT with its receptor CD40.";
 RL Protein Sci. 7:1124-1135(1998).
 [10]
 RN VARIANTS HIGM1 ARG-36 AND GLY-140.
 RX MEDLINE=93156839; PubMed=7679206;
 RA Korthauer U., Graf D., Mages H.W., Briere F., Padayachee M.,
 RA Malcolm S., Ugazio A.G., Notarangelo L.D., Levinsky R.J.,
 RA Kroczeck R.A.;
 RT "Defective expression of T-cell CD40 ligand causes X-linked
 RT immunodeficiency with hyper-IgM.";
 RL Nature 361:539-541(1993).
 [11]
 RN VARIANT HIGM1 GLU-123.
 RX MEDLINE=93156840; PubMed=8094231;
 RA Disanto J.P., Bonnefoy J.Y., Gauchat J.F.M., Fischer A.,

e G.;
utations in X-linked immunodeficiency with hyper-IgM";
-543(1993).

PRO-155; ASP-211 AND VAL-227.
70; PubMed=7679801;
Mitage R.J., Conley M.E., Rosenblatt H., Jenkins N.A.,
Bedell M.A., Edelhoff S., Distche C.M.,
Forslow W.C., Belmont J.W., Spriggs M.K.;
ene defects responsible for X-linked hyper-IgM
0-993(1993).

ALA-126; ARG-140 AND GLU-144.
38; PubMed=7717401;
la A., Strina D., Sacco M.G., Morali F., Brugnoni D.,
ntuano E., Fasth A., Andersson B., Zegers B.J.M.,
znick I., Levy J., Zan-Bar I., Porat Y., Afro P.,
zzoni P., Notarangelo L.D.;
ion of nine novel mutations in the CD40 ligand gene in
X-linked hyper IgM syndrome of various ancestry";
net. 56:898-906(1995).

PRO-155 AND VAL-227, AND VARIANT ARG-219.
33; PubMed=8550833;
J., Allen R.C., Larche M., Greene J.M., Shigeoka A.O.,
rauf D.C., Belmont J.W., Conley M.E.;
nd conformation polymorphism study of CD40 ligand.
tion analysis and carrier detection for X-linked hyper
; 97:196-201(1996).

ARG-36; CYS-140; SER-231; MET-254 AND GLY-227 DEL.
77; PubMed=9150729;
himadzu M., Toru H., Seyama K., Nunoi H., Neubauer M.,
h H.D.;
the CD40 ligand gene in 13 Japanese patients with
-IgM syndrome";
:624-627(1997).

Mediates B-cell proliferation in the absence of co-
s well as IgE production in the presence of IL-4.
n immunoglobulin class switching.
omotrimer.
R LOCATION: Type II membrane protein. Also exists as an
lar soluble form.
CIFICITY: Specifically expressed on activated CD4+
tes.
oluble form derives from the membrane form by

c processing.
effects in TNF α 5 are the cause of X-linked
ciency with hyper-IgM type 1 (HIGM1) [MIM:308230]. HIGM1
noglobulin isotype switch defect characterized by
oncentrations of serum IgM and decreased amounts of all
ypes. Affected males present at an early age (usually
first year of life) recurrent bacterial and
tic infections, including pneumocystis carinii pneumonia
table diarrhea due to cryptosporidium infection. Despite
on treatment with intravenous immunoglobulin, the
agnosis is rather poor, with a death rate of about 10%

Belongs to the tumor necrosis factor family.

NAME=CD40Lbase;

ean CD40L defect database (mutation db);

://www.expasy.org/cd40lbase/";

://ftp.expasy.org/databases/cd40lbase".

NAME=PROW; NOTE=CD guide CD154 entry;

://www.ncbi.nlm.nih.gov/prow/cd/cd154.htm".

OT entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/>
or send an email to license@isb-sib.ch).

CC EMBL; X68550; CAA48554.1; -
DR EMBL; Z15017; CAA78737.1; -
DR EMBL; X67878; CAA48077.1; -
DR EMBL; L07414; AAA35662.1; -
DR EMBL; D31797; BAA06599.1; -
DR EMBL; D31793; BAA06599.1; JOINED.
DR EMBL; D31794; BAA06599.1; JOINED.
DR EMBL; D31795; BAA06599.1; JOINED.
DR EMBL; D31796; BAA06599.1; JOINED.
DR EMBL; D31796; BAA06599.1; JOINED.
DR PIR; S28017; I53476.
DR PDB; 1ALY; 17-SEP-97.
DR PDB; 119R; 22-MAY-02.
DR Genew; HGNC:11935; TNFSF5.
DR MIM; 300386; -
DR MIM; 308230; -
DR GO; GO:0008987; C:integral to plasma membrane; TAS.
DR GO; GO:0005625; C:soluble fraction; TAS.
DR GO; GO:0005174; F:CD40 receptor binding; IPI.
DR GO; GO:0006916; P:anti-apoptosis; IDA.
DR GO; GO:0042100; P:B-cell proliferation; IDA.
DR GO; GO:0006954; P:inflammatory response; IDA.
DR GO; GO:0045190; P:isotype switching; ISS.
DR GO; GO:0007159; P:leukocyte cell adhesion; NAS.
DR GO; GO:0030168; P:platelet activation; IDA.
DR GO; GO:0007165; P:signal transduction; ISS.
DR InterPro; IPR003263; TNF 5.
DR InterPro; IPR006052; TNF family.
DR InterPro; IPR008983; TNF like.

Query Match 7.4%; Score 86; DB 1; Length 261;
Best Local Similarity 25.9%; Pred. No. 2.2;
Matches 35; Conservative 21; Mismatches 55; Indels 24; G
QY 105 TVSGWEE-----TKINSSPLRYDRQIGFTVIRAGLYLYCQVHFDGKAVYLKLDI
DB 136 SVLQWAEKGYTMSNNLVLENGKQL---TVKRQGLYTYAQVTFCSNREASSQAP!
QY 161 GVLAALRCLLEEF-----SATAASSPGQLRLC-----QVSGLLPLRPQSSLRITEL
DB 193 --LCLKSPGRFERILLRAANTHSSAKP-----CQQSIHLGGVFELOFGASVFNVTI
QY 210 LKAAPFLTYFGLFQV 224
DB 247 VSHGTGFTSFGLLKL 261

Search completed: April 7, 2004, 17:45:17
Job time : 10.2829 secs

GenCore version 5.1.6
wright (c) 1993 - 2004 Compugen Ltd.
search, using sw model
7, 2004, 17:41:27 ; Search time 31.3851 Seconds
(without alignments)
2261.954 Million cell updates/sec

245-198A-2
HLGALACLGLLVVVSL.....PWAHLKAAPFLTYFGLFQVH 225
JME2
10.0 , Gapext 0.5
41 seqs, 315518202 residues
satisfying chosen parameters: 1017041
1: 0
1: 2000000000
imum Match 0%
imum Match 100%
ing first 45 summaries

REMBL 25: *
sp archea: *
sp bacteria: *
sp fungi: *
sp human: *
sp invertibrate: *
sp mammal: *
sp mhc: *
sp organelle: *
sp phage: *
sp plant: *
sp rodent: *
sp virus: *
sp vertebrate: *
sp unclassified: *
sp virus: *
sp bacteriap: *
sp archesp: *

he number of results predicted by chance to have a
than or equal to the score of the result being printed,
by analysis of the total score distribution.

SUMMARIES

Y	h	Length	DB	ID	Description
1	410	11	Q8BXS2	Q8BXS2	mus musculus
8	330	4	Q8IZK7	Q8IZK7	homo sapien
0	409	5	Q8IGD3	Q8IGD3	drosophila
5	409	5	Q8MY88	Q8MY88	drosophila
3	261	5	Q8MRW2	Q8MRW2	drosophila
3	325	5	Q8V5G2	Q8V5G2	drosophila
3	415	5	Q8MUJ1	Q8MUJ1	drosophila
9	426	16	Q88IZ6	Q88IZ6	pseudomonas
4	684	13	Q7TJ36	Q7TJ36	lampetra ja
0	557	16	Q8XQX3	Q8XQX3	raistonia s
8	210	16	Q9A926	Q9A926	caulobacter
8	421	16	Q9HUW2	Q9HUW2	pseudomonas
7	213	16	Q82AD2	Q82AD2	streptomyce
7	1363	5	Q9VFD3	Q9VFD3	drosophila
7	287	13	Q90WT9	Q90WT9	gallus gall
6	224	5	Q9V762	Q9V762	drosophila

17	88	7.6	352	12	089341	089341 hendri
18	88	7.6	353	12	Q66760	Q66760 equin
19	88	7.6	532	4	Q16727	Q16727 homo s
20	87.5	7.5	522	10	Q9FTN7	Q9FTN7 oryza
21	87.5	7.5	670	16	Q9AA15	Q9AA15 caulol
22	87	7.5	244	6	Q86227	Q86227 pan tri
23	86.5	7.4	154	16	Q7U7N8	Q7U7N8 syneci
24	86.5	7.4	340	16	Q9HUR8	Q9HUR8 pseud
25	86.5	7.4	504	16	Q92KA4	Q92KA4 rhizol
26	86.5	7.4	1237	11	Q7TPS4	Q7TPS4 mus m
27	86	7.4	260	10	Q8S2N9	Q8S2N9 oryza
28	85.5	7.4	331	10	Q942P9	Q942P9 oryza
29	85.5	7.4	724	5	Q868S9	Q868S9 anophe
30	85.5	7.4	2841	10	Q7XU06	Q7XU06 oryza
31	85	7.3	116	16	Q7V2A2	Q7V2A2 proch
32	84.5	7.3	377	16	Q7UEA6	Q7UEA6 rhodo
33	84.5	7.3	422	16	Q9RKB0	Q9RKB0 strep
34	84.5	7.3	430	2	Q9REU1	Q9REU1 strept
35	84.5	7.3	1079	13	Q8UVR4	Q8UVR4 xenopi
36	84.5	7.3	1118	16	Q98E34	Q98E34 rhizo
37	84	7.2	2962	5	Q93326	Q93326 cacer
38	83.5	7.2	394	16	Q92V66	Q92V66 rhizo
39	83	7.1	467	16	Q9S2Y4	Q9S2Y4 strep
40	82.5	7.1	174	16	Q9CKX1	Q9CKX1 paste
41	82.5	7.1	314	5	Q8WPH7	Q8WPH7 theile
42	82.5	7.1	501	16	Q89IN3	Q89IN3 brady
43	82.5	7.1	549	16	Q8RC38	Q8RC38 therm
44	82.5	7.1	718	6	Q8HXH0	Q8HXH0 macaca
45	82	7.1	619	5	Q8SUH9	Q8SUH9 enceph

ALIGNMENTS

RESULT 1
Q8BXS2 PRELIMINARY; PRT; 410 AA.
AC Q8BXS2; 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Tumor necrosis factor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR ENBL; AK044387; BAC31897.1; -.
DR PIR; PT0714; PT0714.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR006052; TNF_family.
DR SMART; SM00207; TNF; 2.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS00049; TNF_2; 2.
SQ SEQUENCE 410 AA; 45681 MW; 590A4B74C33FB8D4 CRC64;

Query Match 82.1%; Score 954.5; DB 11; Length 410;
Best Local Similarity 88.6%; Pred. No. 1.3e-83;
Matches 195; Conservative 1; Mismatches 13; Indels 11; Gap
1 VLSIGLALCLGILLVVVSLGWSWATLSA-QEPSEELTAEDRPPPLNPQTESQD

GLALACGLLLVSVLSGWTLSAQEPSQEELTAEDRRPELNPQTESQDVVP 84
 LVPRPSAPKGRKAPRAIAAHYVHPRPGDGAQAGVDGTSGWEETKINSASP 119
 LVPRPSAPKGRKAPRAIAAHYVHPRPGDGAQAGVDGTSGWEETKINSASP 144
 RQIGFTVIRAGLYLYCQVHFDEGKAYVLLKDLVNGVLAALRCLEPFSATAASP 179
 RQIGFTVIRAGLYLYCQVHFDEGKAYVLLKDLVNGVLAALRCLEPFSATAASP 204
 RLCOVSGLLPLR-----PGSSLRIRTLPAWHL 210
 RLCOTE-LQSLRREVSRLQRSGPQSGKQGERPQSL 243
 RELIMINARY; PRT; 330 AA.
 TrEMBLrel. 23, Created
 TrEMBLrel. 23, Last sequence update
 TrEMBLrel. 25, Last annotation update
 (Human).
 azoia; Chordata; Craniata; Vertebrata; Euteleostomi;
 eria; Primates; Catarrhini; Hominidae; Homo.
 6;
 N.A.
 24; PubMed12411489;
 B., Medina J.P., Lopez-Fraga M., Lozano J.C.,
 M., Picard A., Martinez-A C., Garcia-Sanz J.A.,
 hybrid mRNA encodes TWE-PRIL, a functional cell surface
 usion protein.";
 1-5720(2002).
 ; AAL90443.1;
 ; C-membrane; IEA.
 ; F.tumor necrosis factor receptor binding; IEA.
 ; P.immune response; IEA.
 06052; TNF family.
 08983; TNF_like.
 TNF; 1.
 ; TNF; 1.
 51; TNF 1; 1.
 49; TNF 2; 2.
 AA; 36588 MW; FC6F3BCA29C029AE CRC64;
 rity 52.8%; Score 613; DB 4; Length 330;
 nservative 8; Mismatches 14; Indels 0; Gaps 0;
 ALACGLLLVSVLSGWTLSAQEPSQEELTAEDRRPELNPQTESQDVVPFL 61
 ALACGLLLVSVLSGWTLSAQEPSQEELVAEEDQPSLNPQTESQDPAFL 85
 PRPSAPKGRKAPRAIAAHYVHPRPGDGAQAGVDGTSGWEETKINSASP 121
 PRPSAPKGRKAPRAIAAHYVHPRPGDGAQAGVDGTSGWEETKINSASP 145
 3EFTVIRAGLYLYCQ 142
 3EFIVTRAGLYLYCQ 166
 ELIMINARY; PRT; 409 AA.
 TrEMBLrel. 23, Created
 TrEMBLrel. 23, Last sequence update
 TrEMBLrel. 25, Last annotation update

GN BCDNA:RH51659.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT001838; AAN71595.1;
 DR FlyBase; FBgn0064801; BCDNA:RH51659.
 DR GO; GO:0016020; C-membrane; IEA.
 DR GO; GO:0005164; F.tumor necrosis factor receptor binding; IEA.
 DR GO; GO:0006955; P.immune response; IEA.
 DR InterPro; IPR006052; TNF_family.
 DR InterPro; IPR008983; TNF_like.
 DR SMART; SM00207; TNF; 1.
 DR PROSITE; PS00251; TNF 1; 1.
 DR PROSITE; PS00049; TNF 2; 1.
 SQ SEQUENCE 409 AA; 46401 MW; FC2E9BD9E012D257 CRC64;
 Query Match 10.0%; Score 116.5; DB 5; Length 409;
 Best Local Similarity 24.3%; Pred. No. 0.01;
 Matches 50; Conservative 32; Mismatches 87; Indels 37; G;
 QY 29 QPSQBELTAEDRRPELNPQTESQDVVPFLQVLRPRRPAKGRKAPRAIA/
 Db 231 QEKSSNEATSKERPAHLHRRMRSH-----RHLLVRKARS-----EDSRP-----AJ
 QY 89 VHRPFGDGAQAGVDGTSGWEETKINSPLRYDRGIGETVIRAGLYLYCQVH
 Db 278 LSSRRRHQSGM-GYHGDYVIGNDNERNYSYQG-HFQTRDGLVTVTNTGLYVVAQIC
 QY 149 KAVYLLKDLVNGVLA-----LRCLEPFSATAASPPQLRLQVSGLLPLRPG
 Db 336 HD-----QNGFTVFGQDTPFLQCLN---TVPTNMPKHVHTCHTSLIHLERNE
 QY 202 IRTL---PWAHLKAPFLTYELFOV 224
 Db 384 LKDIHNDENAVLRGNRNSYFGIFKV 409
 RESULT 4
 Q8MY88 PRELIMINARY; PRT; 409 AA.
 ID Q8MY88
 AC Q8MY88;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE TNF superfamily ligand, Eiger (Tumor necrosis factor family member
 DE DT1).
 GN EIGER OR DT1 OR CG:2919.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22060500; PubMed=12065414;
 RA Igaki T., Kanda H., Yamamoto-Goto Y., Kanuka H., Kuranaga E.,
 RA Aigaki T., Miura M.;
 RT "Eiger, a TNF superfamily ligand that triggers the Drosophila JNK
 RT pathway.";
 RL ENBO J. 21:3009-3018(2002).
 RN [2]

Query Match	9.3%	Score 107.5	DB 5	Length 261
Best Local Similarity	23.1%	Pred. No. 0.042		
Matches	48	Conservative	34	Mismatches 91
				Indels 35
				Gaps 3
QY	29	QEPSEQLTAEDRRPELNPQTEESQDVVFLEQVLRPRS--APGKRKARPRRAIA		
Db	77	QEKSSNEATSKESAPLHRRRMSRH-----RHLLVKGESLISARSDSRP-----A		
QY	87	YEVHPRQDGAQGVDTGVSWEETKINSSPLRYDROIGFTVIRAGLYLYLYCVCH		
QY	128	PHLSRRRHQSGM--GYGDMYIGNDNERNYQG--HFQTRDGLVTNTGLIYYVYALIC		
Db	147	EKGAVYMLDLIVNGVLA-----LRLCEFSATAASPGPQLRQLCVSGLLPLRPG		
QY	166	NSHD-----QNGFIVQGDUTFFLQCLN---TVPTNPHKVHTCHTSGLIHLERN		
QY	200	LRIRTL-----PWAHLKAAFPFLTYGFLQV 224		
Db	234	HLKDIHNDRNAVLREGNRSYFGIFKV 261		
RESULT 6				
Q9V5G2	PRELIMINARY:	PRT:	325	AA.
ID	Q9V5G2			
AC	Q9V5G2			
CD	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	CG12919 protein.			
GN	EIGER OR CG12919.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachyceta; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=Berkeley;			
RX	MEDLINE=20196006; PubMed=107311132;			
RA	Adams M.D., Cainker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers J.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.			
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.			
RA	Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra J.			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hosdin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacleb J.M.,			
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Swirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L			

RA hong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA yers E.W., Rubin G.M., Venter J.C.,
 PR sequence of *Drosophila melanogaster*.;
 BL 185-2195(2000).
 CR 1; AAF58848.1; -.
 CR 0033483; eiger.
 CR 0; C:membrane; IEA.
 CR 4; F:tumor necrosis factor receptor binding; IEA.
 CR 006052; TNF family.
 CR 008983; TNF-like.
 CR 7; TNF; 1.
 CR 251; TNF 1; 1.
 CR 049; TNF 2; 1.
 CR 5 AA; 36862 MW; 685CCB69694FIA3A CRC64;
 CR
 CR 9.3%; Score 107.5; DB 5; Length 325;
 CR arity 23.1%; Pred. No. 0.056;
 CR conservative 34; Mismatches 91; Indels 35; Gaps 9;
 CR
 CR QEELTAEDRREPPPELNQTESQDVVFLQLVPRRS--APKGRKARPRRAAH 86
 CR ||||| : : : : :
 CR SNEATSKESAPLHRRHRSR-----RHLLVRKGESLLSARSDSRP----AAH 191
 CR ||||| : : : : :
 CR PRPQDGAQAGVDGTVSGWETKINSSPLRYDRQIGFTVIRAGLYLYCYVHD 146
 CR ||||| : : : : :
 CR SRRRHQGSQ-GYHGDVYIGNDNERNYSYQG-HFQTRDGLVLTNTGLVYVYVAAICYN 249
 CR ||||| : : : : :
 CR VYLKDLLVNGVLA-----LRCLFEFSATAASSPGQLRLCQVGLPLRPGSS 199
 CR ||||| : : : : :
 CR -----QNGFIVFGQDTFFLQCLN-----TVPTNMPKHVHTCHTSLIHLERNER 297
 CR ||||| : : : : :
 CR TL---PWAHLKAAPFLTYFGLFOV 224
 CR ||||| : : : : :
 CR IHNDRNAVLRGNRRSYFGIFKV 325
 CR ||||| : : : : :
 CR
 CR PRELIMINARY; PRT; 415 AA.
 CR
 CR 'EMBLrel. 22, Created)
 CR 'EMBLrel. 22, Last sequence update)
 CR 'EMBLrel. 25, Last annotation update)
 CR
 CR 19 OR DARTH.
 CR anogaster (fruit fly).
 CR azoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CR pterygota; Diptera; Brachycera; Muscomorpha;
 CR rosophilidae; Drosophila.
 CR 7;
 CR
 CR N.A.
 CR 23; PubMed=12176339;
 CR M., Baer K.;
 CR TNF Signaling Mechanisms. JNK-Dependent Apoptosis
 CR iger, the *Drosophila* Homolog of the TNF Superfamily.";
 CR :1263-1268(2002).
 CR
 CR N.A.
 CR 38; PubMed=12894227;
 CR aaty W.S., Chen P., Tomar R.S., Eby M.T., Chappo J.,
 CR re N., Zachariah S., Sinha S.K., Abrams J.M.,
 CR ; receptor, Wengen, comprise a TNF-like system in
 CR
 CR 60-4867(2003).
 CR ; AAM76710.1; -.
 CR ; AAM66763.1; -.
 CR 033483; eiger.
 CR ; C:membrane; IEA.
 CR ; F:tumor necrosis factor receptor binding; IEA.
 CR ; P:immune response; IEA.

DR InterPro: IPR006052; TNF family.
 DR SMART: IPR008983; TNF_like.
 DR SMART: SM00207; TNF_1.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR SEQUENCE 415 AA; 46918 MW; E087A26DE222DBF CRC64;
 DR
 DR Query Match 9.3%; Score 107.5; DB 5; Length 415;
 DR Best Local Similarity 23.1%; Pred. No. 0.077;
 DR Matches 48; Conservative 34; Mismatches 91; Indels 35; G
 DR
 DR 29 QEPSQEELTAEDRREPPPELNQTESQDVVFLQLVPRRS--APKGRKARPRRA
 DR ||||| : : : : :
 DR 231 QEKSSNEATSKESAPLHRRHRSR-----RHLLVRKGESLLSARSDSRP---
 DR ||||| : : : : :
 DR 87 YEVEHPRPQDGAQAGVDGTVSGWETKINSSPLRYDRQIGFTVIRAGLYLYCYQ
 DR : : : : :
 DR 282 FHLSRRRHQGSQ-GYHGDVYIGNDNERNYSYQG-HFQTRDGLVLTNTGLVYVYVAA
 DR : : : : :
 DR 147 EGRKAVYLKDLLVNGVLA-----LRCLFEFSATAASSPGQLRLCQVGLPLR
 DR ||||| : : : : :
 DR 340 NSHD-----QNGFIVFGQDTFFLQCLN-----TVPTNMPKHVHTCHTSLIHL
 DR : : : : :
 DR 200 LRIRTL---PWAHLKAAPFLTYFGLFOV 224
 DR : : : : :
 DR 388 IHLKDIHNDRNAVLRGNRRSYFGIFKV 415
 DR : : : : :
 DR
 DR RESULT 8
 DR Q881Z6 PRELIMINARY; PRT; 426 AA.
 DR ID Q881Z6
 DR AC Q881Z6;
 DR DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DR DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DR DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DR DE Conserved hypothetical protein.
 DR GN PP2853.
 DR OS Pseudomonas putida (strain KT2440).
 DR OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 DR OC Pseudomonadaceae; Pseudomonas.
 DR OX NCBI_TaxID=160488;
 DR RN [1]
 DR RP SEQUENCE FROM N.A.
 DR RX MEDLINE=2423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes I.
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duisterhoft A., Tuemmler B.,
 RA Fraser C.M.;
 RA "Complete genome sequence and comparative analysis of the
 RT metabolically versatile *Pseudomonas putida* KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AF016784; AAN68461.1; -.
 DR TIGR; PP2853; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 426 AA; 46020 MW; FEDC7E266C982633 CRC64;
 SQ
 SQ Query Match 8.9%; Score 103; DB 16; Length 426;
 SQ Best Local Similarity 29.1%; Pred. No. 0.22;
 SQ Matches 50; Conservative 17; Mismatches 65; Indels 40; Ga
 SQ
 SQ 7 ALACIGLLVVVSLGSAWATLSAQEPSQEELTAEDRREPPPELNQTESQDVVFLQ
 SQ ||||| : : : : :
 SQ 9 AFVCLTTTLAPATAL-----YAADPQVEAL-----RQELIELKRRYEAQQALMWLEQ
 SQ ||||| : : : : :
 SQ 67 PRRSAPKGRKARPRRAIAAHVEVHPRPQDGAQ-----AGVDGTV-SGWETKINSSP
 SQ ||||| : : : : :
 SQ 60 QVEEAPAA--AQPKRLVKS-----PAGVKGAQTVASGAPGTGGTSSYQALTAIDSEP
 SQ ||||| : : : : :
 SQ

79; PubMed=11823852;
Genin S., Artiguenave F., Guzy J., Manganot S.,
ault A., Brotier P., Camus J.C., Catcolico L.,
oisne N., Claudel-Renard C., Cunnac S., Demange N.,
le M., Moisan A., Robert C., Saurin W., Schlex T.,
ebault P., Whalen M., Winkler P., Levy M.,
Boucher C.A.;
ce of the plant pathogen *Ralstonia solanacearum*.
502(2002).

; B41; 1.
 ; PDZ; 1.
 ; WW; 1.
 57; FERM_3; 1.
 56; PDZ; 1.
 59; WW DOMAIN; 1.

Query Match	7.7%;	Score 89;	DB 13;	Length 287;
Best Local Similarity	23.4%;	Pred. No. 2.9;		
Matches	40;	Conservative	24;	Mismatches 65; Indels 42; Gaps 4
QY	84	AAHYEVHPR-PQCGAQAQVDG-----TVSGWEETKINSS-SPLRYDRQIGBEFTV		
DB	127	SAHLFRCPNPAQDSSRRFGNLSQSCHEAITRWEDSTIHSHLQNTVY--RDGRLRV		
QY	135	GLYYLYCQVHF-----EGKAVYIKLDDLLVNGVLAALRCLBEFS		
DB	185	GKYYVYSQIYFRYSDGAGRVPLVQCINWTSYSQPIILLKGV-----		
QY	176	ASSPGPQ--LRLCQVSGLLPLRPGSSLRIRTLPAHLKAAPFLTYFGLFQV 224		

06:25:13 2004

us-09-245-198a-2.rspt

EAEGHLYQGLFELKAGDELFSVSSLAIDYSDAAASYEGAFL 285

April 7, 2004, 17:46:43
secs

16:25:12 2004

us-09-245-198a-2.rag

1

GenCore version 5.1.6
copyright (c) 1993 - 2004 Compugen Ltd.

n search, using sw model

il 7, 2004, 17:37:32 ; Search time 45.0884 Seconds
(without alignments)
1409.967 Million cell updates/sec

09-245-198A-2
2
LSLGLALACTLLLVVWSL.....PWAHLKARPELTFFGLFQW 225

SUM62

op 10.0 , Gapext 0.5

6107 segs, 282547505 residues

s satisfying chosen parameters: 1586107

th: 0
th: 2000000000

nimum Match 0%
ximum Match 100%
string first 45 summaries

Genesecp_29Jan04: *
genesecp1980s: *
genesecp1990s: *
genesecp2000s: *
genesecp2001s: *
genesecp2002s: *
genesecp2003as: *
genesecp2003bs: *
genesecp2004s: *

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
id by analysis of the total score distribution.

SUMMARIES

ry	ch	Length	DB	ID	Description
1.0	225	2	AAW47524		Aaw47524 Mus muscu
1.0	225	3	AAB07527		Aab07527 Amino aci
1.0	249	7	ADC97712		Adc97712 Murine FL
1.7	211	2	AAW93591		Aaw93591 Mouse TNF
1.8	249	2	AAW29745		Aaw29745 TNF relat
1.8	249	2	AAU09369		Aau09369 Human tum
1.8	249	3	AAU95338		Aau95338 Human PRO
1.8	249	3	AAB07526		Aab07526 Amino aci
1.8	249	4	AAE00891		Aae00891 Human TRE
1.8	249	5	AAU986129		Aau986129 Human PRO
1.8	249	6	ABR42315		AbR42315 Human TRE
1.8	249	7	ADC35206		Adc35206 Human TNF
1.8	284	2	AAW47525		Aaw47525 Homo sapi
1.8	208	2	AAW93590		Aaw93590 Human TNF
1.8	273	4	AAU03499		Aau03499 TWEAK ext
1.9	189	2	AAW29746		Aaw29746 TNF relat
1.9	189	4	AAE00892		Aae00892 Human UL4
1.9	146	4	AAE00895		Aae00895 Human TRE
1.6	406	5	AAU77717		Aau77717 Drosophil
1.5	409	5	AAU77718		Aau77718 Drosophil
1.3	325	4	ABB67553		Abb67553 Drosophil
1.2	294	2	AAW69956		Aaw69956 NF-kB rec
1.2	294	2	AAW68292		Aaw68292 NF-kB rec
1.2	294	2	AAE08737		Aae08737 Murine re
1.2	294	4	AAE04425		Aae04425 Murine re

26	106.5	9.2	294	4	AAE01992	Mu
27	106.5	9.2	294	5	AAE26102	Mol
28	106.5	9.2	294	7	ADB16986	Mu
29	106.5	9.2	294	7	ADC73000	Mu
30	106.5	9.2	294	7	ADC78266	Mur
31	103	8.9	220	4	AAE62340	GpJ
32	103	8.9	426	6	ABU39962	Prc
33	102.5	8.8	316	2	AAW83017	Ost
34	102.5	8.8	316	2	AAW83194	Hun
35	102.5	8.8	316	2	AAW59654	Ami
36	102.5	8.8	316	2	AAI17874	Mur
37	102.5	8.8	316	3	AAU91024	Mol
38	102.5	8.8	316	3	AAU84418	Ami
39	102.5	8.8	316	3	AAU84419	Ami
40	102.5	8.8	316	5	AAU78289	Mol
41	102.5	8.8	316	6	ABR42071	Hun
42	102.5	8.8	316	6	ABR99477	Ami
43	102.5	8.8	316	6	ABU08463	Ami
44	102.5	8.8	316	6	ABR55560	Ami
45	99	8.5	234	4	AAE62339	GpJ

ALIGNMENTS

RESULT 1	
AAW47524	AAW47524 standard; protein; 225 AA.
XX	AC AAW47524;
XX	AC
XX	21-JUL-1998 (first entry)
XX	DE Mus musculus tumour necrosis factor related ligand (TRELL).
XX	XX TRELL; tumour necrosis factor related ligand; tnfr; treatment; can
KW	autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection.
XX	XX Mus musculus.
OS	XX Mus musculus.
XX	XX
XX	Key Location/Qualifiers
FT	Domain 1..21
FT	/note= "hydrophobic, transmembrane domain"
XX	XX
XX	WO9805783-A1.
XX	XX
PD	12-FEB-1998.
XX	XX
XX	07-AUG-1997; 57WO-US013945.
XX	XX
PR	07-AUG-1996; 96US-0023541P.
PR	18-OCT-1996; 96US-0028515P.
PR	18-MAR-1997; 97US-0040820P.
XX	XX
PA	(BIOJ) BIOGEN INC.
XX	(UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX	XX
PI	Chicheportiche Y, Browning JL;
XX	XX
DR	WPI; 1998-145619/13.
DR	N-PSDB; AAV18599.
XX	XX
PT	Tumour necrosis factor related ligand - useful for, e.g. treating
PT	auto-immune disease and immune responses to tissue grafts.
XX	XX
PS	Claim 12; Page 48-50; 69pp; English.
XX	XX
CC	The sequence is that of mouse tumour necrosis factor related liga
CC	(TRELL). TRELL or active fragments can be included with a carrier
CC	pharmaceutical compositions to treat cancer, autoimmune diseases,
CC	immune responses to tissue grafts, or to stimulate or suppress th
CC	system. It is useful to screen for TRELL receptors, by labelling

nistering to the subject a TWEAK agonist or antagonist.

ID NO 1; 120pp; English.

quence is murine transmembrane FL-TWEAK (TNF relatedness
ty to induce cell death, where TNF is Tumour Necrosis
is a member of the TNF family. TWEAK agonists or
e useful for treating a TWEAK-related condition, e.g.
iac disease; liver disease; lung disease; kidney disease;
skeletal muscle disease; adipose tissue disease;
al tract disease; pancreatic disease; reproductive organ
1 disease; cartilage disease; bone disease; connective
; cellular death; and a pathological condition of a tissue
WEAK receptor.

A;

100.0%; Score 1162; DB 7; Length 249;
rity 100.0%; Pred. No. 3.6e-112;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

LALACLGLLVVSLGSWATLSAQEPSOEELTAEDRREPELNPQTEESQDVVFP 60
|||||
LALACLGLLVVSLGSWATLSAQEPSOEELTAEDRREPELNPQTEESQDVVFP 84

RPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGTGSGWEETKINSSPL 120
|||||
RPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGTGSGWEETKINSSPL 144

IGFTVIRAGLYLYCYQVHFDEGKAVYIKLDLLVNGVLALRCLBEFSATASSPG 180
|||||
IGFTVIRAGLYLYCYQVHFDEGKAVYIKLDLLVNGVLALRCLBEFSATASSPG 204

COVSGLLPLRPGSSLRIRTLPAHLKAAPFLTYFGLFQVH 225
|||||
COVSGLLPLRPGSSLRIRTLPAHLKAAPFLTYFGLFQVH 249

lard; protein; 211 AA.

(first entry)

otein.

s factor receptor; signal transducer molecule; TNF; APO4;
abnormality; gestational abnormality; prostate cancer;
PO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
main; immunogen; antibody preparation; breast carcinoma;
ise.

98WO-US018393.

97US-00924634.

WASHINGTON.

191/17.

425.

rosis Factor family receptor polypeptides and ligands -
agnosis and treatment of prostate cancer and developmental

or gestational abnormalities.

Claim 40; Fig 13B; 156pp; English.

This invention describes isolated Tumor Necrosis Factor (TNF) fami
receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNF
their active fragments. APO4 is useful for diagnosing prostate car
determining levels of APO4 in an individual. Prostate cancer can
treated using APO4 selective binding agents linked to a therapeutic
moisty. APO4 polypeptides are also useful for identifying selectiv
binding agents, useful in diagnosis/treatment of disease by bindin
agents to the polypeptide/active fragment which is extracellular,
expressed on the cell surface. The binding is preferably performe
vivo. APO4 polypeptides/ active fragments are also useful for scr
for agonists and antagonists by binding and observing the changer
activity. Effective pharmacological agents useful in diagnosis or
treatment of disease are also identified using APO4 polypeptides/
fragments and APO4 signal transducer molecules that specifically
with a cytoplasmic domain of APO4 and detecting a change in level
activity. The method is performed in vivo or in vitro. APO polypep
are all useful as immunogens for preparing antibodies. APO4 is al
useful for diagnosis/treatment of developmental or gestational
abnormalities. APO8 was transfected to human breast carcinoma cel.
MCF-7, and induced apoptosis

Sequence 211 AA;

Query Match 93.7%; Score 1089; DB 2; Length 211;
Best Local Similarity 99.5%; Pred. No. 1.1e-104;
Matches 210; Conservative 0; Mismatches 1; Indels 0; G

QY 15 LVVSLGSWATLSAQEPSOEELTAEDRREPELNPQTEESQDVVPEQLVPRRS;
Db 1 LVVSLGSWATLSAQEPSOEELTAEDRREPELNPQTEESQDVVPEQLVPRRS;

QY 75 RKARPRRAIAAHYEVHPRPGDGAQAGVDGTGSGWEETKINSSPLRYDROIGFT
Db 61 RKARPRRAIAAHYEVHPRPGDGAQAGVDGTGSGWEETKINSSPLRYDROIGFT

QY 135 GLYLYCYQVHFDEGKAVYIKLDLLVNGVLALRCLBEFSATASSPGRLRCQVSG
Db 121 GLYLYCYQVHFDEGKAVYIKLDLLVNGVLALRCLBEFSATASSPGRLRCQVSG

QY 195 RPSGSSLRIRTLPAHLKAAPFLTYFGLFQVH 225

Db 181 RPSGSSLRIRTLPAHLKAAPFLTYFGLFQVH 211

RESULT 5
AAW29745
ID AAW29745 standard; protein; 249 AA.

XX AC AAW29745;

XX DT 27-OCT-1998 (first entry)

XX DE TNF related endothelium proliferative agent protein.

XX KW TNF; endothelium proliferative agent; TREPA; wound healing; cance

XX KW tissue grafting; vascularisation; apoptosis; autoimmune; birth co

XX OS Homo sapiens.

XX PN WO9835061-A2.

XX PD 13-AUG-1998.

XX PF 12-FEB-1998; 98WO-US002859.

XX PR 12-FEB-1997; 97US-00798692.

XX PR 10-FEB-1998; 98US-00021706.

XX

IGEFIVIRAGLYLYCQVHFDEGKAVYKLDLLVDGVLAURCLBEFSATAASSIGP 205
 CQVSGLLPLRPGSSLRIRTLTPWAHLKAAPFLTYFGLFQVH 225
 CQVSGLLALRPGSSLRIRTLTPWAHLKAAPFLTYFGLFQVH 249

dard; protein; 249 AA.

(first entry)

TNF related endothelium proliferative agent).

necrosis factor; TNF; angiogenesis; wound healing; TREPA;
 endothelium proliferative agent; tumour; metastasis;
 nerary.

Location/Qualifiers

98..249

/label= Extracellular_domain

98US-00105343.

97US-00798692.

98US-00021706.

T LAB.

760/29.

350.

ogenesis in mammal at desired sites for promoting wound
 administering soluble fragment of extracellular domain of
 s factor related endothelium proliferative agent protein.

75-76; 53pp; English.

vention relates to extracellular signal molecules,
 members of tumour necrosis factor (TNF) family molecules
 TREPA (TNF related endothelium proliferative agent).
 gically active TREPA are used to treat TREPA-associated
 ours or metastases. TREPA is used for inducing angiogenesis
 promoting wound healing and for vascularising grafted tissue
 l grafting and to promote tissue grafts. The present amino
 is clone ID #690050 human TREPA

AA;

arity 87.8%; Score 1020; DB 4; Length 249;

conservative 9; Mismatches 16; Indels 0; Gaps 0;

HALACLGLLVVSLGSWATLSAQPSQBELTAEDRREPPPELNPOTESSQDVPFL 61

HALACLGLLVVSLGSPASLSAQEPAPAEELVAEEDQPSPELNPOTESSQDPAFL 85

TPRRSAPKGRKARPRRAIAAHYVHPRPQDGAQGVDTGTVSGWEETKINSSPLR 121

TPRRSAPKGRKTRARRAIAAHYVHPRPQDGAQGVDTGTVSGWEETKINSSPLR 145

IGEFIVIRAGLYLYCQVHFDEGKAVYKLDLLVNGVLAURCLBEFSATAASSFGP 181

Db 146 YNRQIGFIVIRAGLYLYCQVHFDEGKAVYKLDLLVDGVLAURCLBEFSATAAS
 Qy 182 QLRCCQVSGLLPLRPGSSLRIRTLTPWAHLKAAPFLTYFGLFQVH 225
 Db 206 QLRCCQVSGLLALRPGSSLRIRTLTPWAHLKAAPFLTYFGLFQVH 249

RESULT 10

AAU86129

ID AAU86129 standard; protein; 249 AA.

XX AC AAU86129;

DT 15-JUL-2002 (first entry)

DE Human PRO207 polypeptide.

XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
 KW leukaemia; neuronal disorder; stromal disorder; blastocoeleic disc
 KW inflammatory disorder; immune disorder; angiogenic disorder; cyt
 KW neuroprotective.

XX OS Homo sapiens.

XX FN WO200153486-A1.

XX PD 26-JUL-2001.

XX PF 11-FEB-2000; 2000WO-US003565.

XX PR 08-MAR-1999; 99WO-US005028.

XX PR 11-MAR-1999; 99US-0123972P.

XX PR 11-MAY-1999; 99US-0133459P.

XX PR 02-JUN-1999; 99WO-US012252.

XX PR 22-JUN-1999; 99US-0140650P.

XX PR 22-JUN-1999; 99US-0140653P.

XX PR 20-JUL-1999; 99US-0144758P.

XX PR 26-JUL-1999; 99US-0145698P.

XX PR 17-AUG-1999; 99US-0146222P.

XX PR 31-AUG-1999; 99US-0149395P.

XX PR 01-SEP-1999; 99US-0151689P.

XX PR 15-SEP-1999; 99WO-US021090.

XX PR 30-NOV-1999; 99WO-US028313.

XX PR 01-DEC-1999; 99WO-US028301.

XX PR 01-DEC-1999; 99WO-US028634.

XX PR 05-JAN-2000; 2000WO-US000219.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
 Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
 Watanabe CK, Wood WI;

XX WPI; 2002-205567/26.

XX N-PSDB; ABK40255.

XX Thirty five nucleic acids encoding PRO polypeptides, useful for t
 benign or malignant tumors, leukemias and lymphoid malignancies,
 inflammatory, angiogenic and immunologic disorders.

XX Claim 61; Fig 4; 302pp; English.

XX The present invention relates to the isolation of novel human PRO
 polypeptides and the polynucleotide sequences encoding them. The
 polypeptides, agonists, antagonists or anti-PRO antibodies are us
 treating benign or malignant tumours (e.g. renal, kidney, bladder;
 breast, etc), leukaemias and lymphoid malignancies, other disorde
 as neuronal, glial, astrocytal, hypothalamic, glandular, macroph
 stromal and blastocoeleic disorders, inflammatory, immune and angi
 disorders. The polynucleotide sequences are also useful in gene t
 CC AAU86128-AAU86162 represent the human PRO polypeptides of the inv

A; 87.8%; Score 1020; DB 5; Length 249;
 rity 88.8%; Pred. No. 2.1e-97;
 nservative 9; Mismatches 16; Indels 0; Gaps 0;
 ALACLGILLVVVSLGSWATLSAQEPSQBELTAEDRRPEELNPQTESQDVPEL 61
 ALACLGILLVAVSLGSRASLSAQEPQAQELVAEEDQDPSELNPQTESQDPAFL 85
 PRRAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSSPLR 121
 PRRAPKGRKTRARRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSSPLR 145
 GEFTVIRAGLYLYCQVHFDGKAVYLLKDLLVNGVLALRCLEBFSATAASS 181
 GEFIVTRAGLYLYCQVHFDGKAVYLLKDLLVNGVLALRCLEBFSATAASSLP 205
 QVSGLLPLRPGSSLRIRTLPAWHLKAAPFLTYFGLFQVH 225
 QVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYFGLFQVH 249

(ard; protein; 249 AA.

first entry)

otein.

tumour necrosis factor; ligand; cytostatic;
 r; osteopathic.

2.

002WO-US023782.

001US-0307838P.

GENOME SCI INC.

osen CA;

59/40.

01.

imeric complex having a first polypeptide member of the
 factor (TNF) ligand family, and a second different member
 family, useful for treating cancer, osteoporosis or an
 ease.

ge 368-369; 388pp; English.

quence is the protein sequence for human TNFAIP3 protein. The
 tes to compositions comprising heterotrimeric complexes of
 s factor (TNF) ligand family members, and their use in the
 vention and treatment of disease. In one embodiment, the
 : complex comprises full-length or extracellular portions of
 -length or extracellular portions of other TNF ligand
 ; preferably VEGI or VEGI-SV. The heterotrimeric complexes
 on are useful for treating an autoimmune disease, cancer or
 and particularly for inhibiting cancer cell proliferation,
 cell proliferation, or inducing apoptosis of T cells

A;

Query Match 87.8%; Score 1020; DB 6; Length 249;
 Best Local Similarity 88.8%; Pred. No. 2.1e-97;
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;
 QY 2 LSLGIALACLGILLVVVSLGSWATLSAQEPSQBELTAEDRRPEELNPQTESQDV 1
 Db 26 LGIGLALACLGILLVAVSLGSRASLSAQEPQAQELVAEEDQDPSELNPQTESQD 1
 QY 62 EOLVPRPSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSS 1
 Db 86 NKLIVPRPSAPKGRKTRARRAIAAHYEVHPRPGDGAQAGVDTVSGWEETKINSS 1
 QY 122 YDRQIGEFIVTRAGLYLYCQVHFDGKAVYLLKDLLVNGVLALRCLEBFSATAASS 1
 Db 146 YNRQIGEFIVTRAGLYLYCQVHFDGKAVYLLKDLLVNGVLALRCLEBFSATAASS 1
 QY 182 QLRQVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYFGLFQVH 225
 Db 206 QLRQVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYFGLFQVH 249

RESULT 12

ADC35206

ID ADC35206 standard; protein; 249 AA.

XX ADC35206;

XX 18-DEC-2003 (first entry)

XX Human TNF ligand family member #12.

XX human; tumour necrosis factor; TNF ligand; endokine alpha;
 KW excessive bone resorption disorder; osteoporosis; Paget's disease.
 KW arterial calcification.

XX Homo sapiens.

XX US2003100074-A1.

XX 29-MAY-2003.

XX 15-AUG-2002; 2002US-00218547.

XX 16-AUG-2001; 2001US-0312542P.

PR 30-OCT-2001; 2001US-0330761P.

XX (YUGG/) YU G.

PA (NIJU/) NI J.

PA (ROSE/) ROSEN C A.

PA (NARD/) NARDELLI B.

XX Yu G, Ni J, Rosen CA, Nardelli B;

XX WPI; 2003-696072/66.

DR N-PSDB; ADC35205.

XX New Endokine alpha gene useful for preparing a composition for treat-
 disease associated with excessive or insufficient bone resorption
 PT osteoporosis, Paget's disease or arterial calcification.

XX Disclosure; SEQ ID NO 24; 145pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding
 CC tumour necrosis factor family ligand. A composition comprising the
 CC isolated antibody or its fragment is used for treating an individual
 CC need of decreased level of endokine alpha activity. The endokine
 CC polypeptide present in a heterotrimeric complex is used for treat-
 CC individual having a disorder associated with excessive bone resor-
 CC e.g. osteoporosis, Paget's disease or arterial calcification. Tre-
 CC individual having a disorder associated with insufficient bone res-
 CC comprises administering an endokine alpha antagonist, which is the
 CC antibody that binds specifically to endokine alpha polypeptide. T-

ence represents the amino acid sequence of a tumour necrosis factor related ligand.

AA;

Identity 87.8%; Score 1020; DB 7; Length 249;
Conservative 9; Mismatches 16; Indels 0; Gaps 0;

HLAALCLGLLVVSLGSGWATLSAQEPSCBELTAEDRRPELNPQTESQDVVPEL 61

HLAALCLGLLVVSLGSGWATLSAQEPSCBELTAEDRRPELNPQTESQDVVPEL 85

/PRRSAPKGRKARRPRAIAAHYEVHPRPGDGAQAGVDGTVSGWEETKINSPLR 121

/PRRSAPKGRKARRPRAIAAHYEVHPRPGDGAQAGVDGTVSGWEETKINSPLR 145

/IGFTVIRAGLYLYLCVHFDEGKAVYKLDLLVNGVLALRCLFEFSATASSPPG 181

/IGFTVIRAGLYLYLCVHFDEGKAVYKLDLLVNGVLALRCLFEFSATASSLP 205

/QVSGLLPLRPSSLRIRTLPAHLKAAPFLTYFGLFOVH 225

/QVSGLLPLRPSSLRIRTLPAHLKAAPFLTYFGLFOVH 249

Standard; protein; 284 AA.

(first entry)

tumour necrosis factor related ligand (TRELL).

; necrosis factor related ligand; tnfr; treatment; cancer;
; disease; immune system; stimulation; suppression;
; ON.

97WO-US013945.

96US-0023541P.

96US-0028515P.

97US-0040820P.

IN INC.
GENEVA FACULTY MEDICINE.

ie Y, Browning JL;

619/13.

8600.

is factor related ligand - useful for, e.g. treating cancer,
disease and immune responses to tissue grafts.

je 50-51; 69pp; English.

is that of human tumour necrosis factor related ligand
JL or active fragments can be included with a carrier in
al compositions to treat cancer, autoimmune diseases or
uses to tissue grafts, or to stimulate or suppress the immune
; useful to screen for TRELL receptors, by labelling with a
abel and screening compositions for binding. Agents
vich TRELL-receptor binding can also be screened for, can
listered, optionally with interferon- gamma, to induce cell

CC death or treat, suppress or alter immune responses (especially in
CC human adenocarcinoma cells) involving a signal pathway between T
CC its receptor. It's coding sequence can be used in gene therapy f
CC related disorders in mammals (especially humans), e.g. tumours,
CC autoimmune and inflammatory diseases or inherited genetic disord
CC introducing into cells, and expressing, therapeutically effectiv
CC of a vector, e.g. a virus comprising a gene encoding TRELL. It m
CC be of use in the preparation of prepare probes for screening
CC natural/synthetic DNAs for TRELL-encoding sequences and for anti
CC therapy

XX Sequence 284 AA;

Query Match 87.8%; Score 1020; DB 2; Length 284;
Best Local Similarity 88.8%; Pred. No. 2.6e-97;
Matches 199; Conservative 9; Mismatches 16; Indels 0;

QY 2 LSLGLALACLGILLVAVSLGSGWATLSAQEPSCBELTAEDRRPELNPQTESQD

Db 61 LSLGLALACLGILLVAVSLGSGWATLSAQEPSCBELTAEDRRPELNPQTESQD

QY 62 SOLVPRPSAPKGRKARRPRAIAAHYEVHPRPGDGAQAGVDGTVSGWEETKINS

Db 121 NRVPRPSAPKGRKARRPRAIAAHYEVHPRPGDGAQAGVDGTVSGWEETKINS

QY 122 YDQIGFTVIRAGLYLYLCVHFDEGKAVYKLDLLVNGVLALRCLFEFSATAA

Db 181 YNRQIGFTVIRAGLYLYLCVHFDEGKAVYKLDLLVNGVLALRCLFEFSATAA

QY 182 QLRLCQVSGLLPLRPSSLRIRTLPAHLKAAPFLTYFGLFOVH 225

Db 241 QLRLCQVSGLLPLRPSSLRIRTLPAHLKAAPFLTYFGLFOVH 284

RESULT 14

AAW93590

ID AAW93590 standard; protein; 208 AA.

XX AC AAW93590;

DT 18-JUN-1999 (first entry)

DE Human TNRL3 protein.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF
KW developmental abnormality; Gestational abnormality; prostate c
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW cytoplasmic domain; immunogen; antibody preparation; breast carc
XX apoptosis; human.

OS Homo sapiens.

PN WO9911791-A2.

XX PD 11-MAR-1999.

XX PF 04-SEP-1998; 98WO-US018393.

XX PR 05-SEP-1997; 97US-00924634.

XX (UNIW) UNIV WASHINGTON.

XX PI Chaudhary PM;

XX DR WPI; 1999-205191/17.

XX DR N-PSDB; AAX23424.

XX New Tumor Necrosis Factor family receptor polypeptides and ligand
PT useful for diagnosis and treatment of prostate cancer and develop
PT or gestational abnormalities.

XX Claim 40; Fig 13A; 156pp; English.

XX

describes isolated Tumor Necrosis Factor (TNF) family peptides: APO4, APO6, APO8 and APO9 or their active isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or fragments. APO4 is useful for diagnosing prostate cancer by levels of APO4 in an individual. Prostate cancer can also be APO4 selective binding agents linked to a therapeutic peptide. APO4 is also useful for identifying selective peptide/active fragment of disease by binding of he cell surface. The binding is preferably performed in peptides/ active fragments are also useful for screening and antagonists by binding and observing the change in APO4 active pharmacological agents useful in diagnosis or disease are also identified using APO4 polypeptides/active APO4 signal transducer molecules that specifically interact with domain of APO4 and detecting a change in level of APO4 method is performed in vivo or in vitro. APO polypeptides as immunogens for preparing antibodies. APO4 is also used for treatment of developmental or gestational APO8 was transfected to human breast carcinoma cell line used apoptosis

81.8%; Score 951; DB 2; Length 208;
rity 88.9%; Pred. No. 2.5e-90;
nservative 8; Mismatches 15; Indels 0; Gaps 0;
WATLSAQEPSEELTAEDRPPPELNPQTEESQDVVFLQVLRPSAPKGRKA 77
RASLSAQEPSEELVAEEDQPSSELPQTEESQDPAPFLNRLVPRSPAPKGRKT 60
IAAHYEVHPRPGQAGAGVDGTVSGWEETKINSPLRYDROQIGFTVIRAGLY 137
IAAHYEVHPRPGQAGAGVDGTVSGWEERINSSPLRYQIGFTVIRAGLY 120
VHFDEGKAVYKLDLLVNGVIALRCLEEFSAATASSPGPQLRCQVSGLLPLRP 197
VHFDEGKAVYKLDLLVNGVIALRCLEEFSAATASSPGPQLRCQVSGLLALRPG 180
RTLFWAHLKAAPFLTYGLFQVH 225
RTLFWAHLKAAPFLTYGLFQVH 208

ard; protein; 273 AA.

first entry)

lular domain-containing fusion protein.

lular domain; tumour necrosis factor; TNF; angiogenesis; ularisation; diabetic retinopathy; neovascular glaucoma; ; retinopathy of prematurity; retrolental fibroplasia; tis; macular degeneration; arthritis; rheumatism; neovascularisation; psoriasis; metastatic condition; ur; sarcoma; carcinoma; benign tumour; haemophilic joint; condition; myocardial angiogenesis; wound granulation; ascular adhesion; telangiectasia; ischaemia; human; c plaque neovascularisation; coronary atherosclerosis; erosclerosis; PDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;

XX
PF
XX
PR
XX
PA
XX
PI
XX
DR
DR
XX
PT
PT
PT
XX
PS
XX

19-DEC-2000; 2000WO-US034755.

20-DEC-1999; 99US-0172878P.

10-MAY-2000; 2000US-0203347P.

(IMMV) IMMUNEX CORP.

Wiley SR;

WPI: 2001-417975/44.

N-PSDB; AAS03964.

Modulating angiogenesis in a mammal for treating diseases mediated angiogenesis, e.g. solid tumors and vascular deficiencies of cardi peripheral tissue, by administering antagonist or agonist of TWEAK receptor.

Example 1; Page 41; 46pp; English.

The sequence represents a fusion protein encoded by the expressi vector pDC409-LZ-TWEAK. The fusion protein comprises a growth horn leader, a leucine zipper multimerisation domain, and the extracell domain of human TWEAK. The fusion protein was used in the isolatic human TWEAK receptor (TWEAKR)-expressing clones from a COS cell in cDNA library. The TWEAK protein is a member of the tumour necrosis (TNF) family and induces angiogenesis. TWEAKR may therefore be use screen for and develop TWEAKR agonists and antagonists for the mod of angiogenesis, to be used in the treatment and diagnosis of hme disease. The disorders mediated by angiogenesis include ocular dis characterised by ocular neovascularisation such as diabetic retin neovascular glaucoma, retinoblastoma, retinopathy of prematurity, retrolental fibroplasia, rubeosis, uveitis, macular degeneration, corneal graft neovascularisation, and inflammatory diseases such a arthritis, rheumatism and psoriasis. Other treatable diseases incl malignant and metastatic conditions such as sarcomas and carcinoma benign tumours and preneoplastic conditions, myocardial angiogenes haemophilic joints, scleroderma, vascular adhesions, atherosclerot plaque neovascularisation, telangiectasia, wound granulation, corc atherosclerosis, peripheral atherosclerosis and ischaemia

Sequence 273 AA;

Query Match 81.8%; Score 951; DB 4; Length 273;
Best Local Similarity 88.9%; Pred. No. 3.6e-90;
Matches 184; Conservative 9; Mismatches 14; Indels 0; G

QY 19 SLGSWATLSAQEPSEELTAEDRPPPELNPQTEESQDVVFLQVLRPSAPKGR
DB 67 SLGSRASLSAQEPSEELVAEEDQPSSELPQTEESQDPAPFLNRLVPRSPAPKGR

QY 79 PRAIAAHYEVHPRPGQAGAGVDGTVSGWEETKINSPLRYDROQIGFTVIRAC
DB 127 ARRAIAAHYEVHPRPGQAGAGVDGTVSGWEERINSSPLRYQIGFTVIRAC

QY 139 LYCQVHFDEGKAVYKLDLLVNGVIALRCLEEFSAATASSPGPQLRCQVSGLLPLF
DB 187 LYCQVHFDEGKAVYKLDLLVNGVIALRCLEEFSAATASSPGPQLRCQVSGLLALF

QY 199 SIRIRTLFWAHLKAAPFLTYGLFQVH 225

DB 247 SIRIRTLFWAHLKAAPFLTYGLFQVH 273

Search completed: April 7, 2004, 17:44:45
Job time : 48.0884 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

ic search, using sw model

ril 7, 2004, 17:30:19 ; Search time 3328.48 Seconds
(without alignments)
12318.149 Million cell updates/sec

-09-245-198a-3

73

atgcattgttagacttga.....gacaaaatgtgataaatgg 1373

ENTITY NUC

pop 10.0 , Gapext 1.0

513289 segs, 14931090276 residues

ts satisfying chosen parameters: 55026578

gth: 0

gth: 2000000000

inimum Match 0%

aximum Match 100%

isting first 45 summaries

ST.*

em_estba:*

em_esthum:*

em_estin:*

em_estmu:*

em_estov:*

em_estpl:*

em_estro:*

em_htc:*

gb_est1:*

gb_est2:*

gb_hic:*

gb_est3:*

gb_est4:*

gb_est5:*

em_estfun:*

em_estom:*

em_gss_hum:*

em_gss_inv:*

em_gss_pln:*

em_gss_vrt:*

em_gss_fun:*

em_gss_mam:*

em_gss_mus:*

em_gss_pro:*

em_gss_rtd:*

em_gss_png:*

em_gss_vrl:*

gb_gss1:*

gb_gss2:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
ad by analysis of the total score distribution.

SUMMARIES

rch	Length	DB	ID	Description
1.2	776	13	BX090012	BX090012
1.8	834	12	BI766766	BI766766 603056866
1.9	948	13	BQ707185	BQ707185 AGENCOURT
1.4	963	13	BQ671259	BQ671259 AGENCOURT

5	743.4	54.1	777	12	BI19200	6
6	725.4	52.8	1071	12	BM921213	7
7	701.4	51.1	731	12	BI871711	8
8	688	50.1	828	12	BI596681	9
9	677.4	49.3	728	12	BI870393	10
10	660.4	48.1	697	14	CF126932	11
11	656.6	47.8	666	14	CF126539	12
12	654.4	47.6	609	12	BI966060	13
13	581.4	42.3	1027	28	AF163779	14
14	567	41.3	567	14	CA396679	15
15	562	40.9	940	13	BQ884231	16
16	560.6	40.8	2237	11	AK044387	17
17	545	39.7	545	14	CB141389	18
18	544.4	39.7	575	14	CA413067	19
19	538.8	39.2	569	14	CB529199	20
20	535.6	39.0	569	13	BU631264	21
21	534.6	38.9	568	12	BM971606	22
22	531.6	38.7	910	12	BG110063	23
23	516.8	37.6	531	12	BI824443	24
24	494.2	36.0	824	14	CB998034	25
25	493.6	36.0	918	10	BF577781	26
26	462	33.6	474	13	BU951915	27
27	458.2	33.4	1033	11	AK020909	28
28	454.4	33.1	561	10	AW763237	29
29	444	32.3	951	13	BQ674188	30
30	442.8	32.3	587	12	BG686319	31
31	440.2	32.1	456	12	BI966255	32
32	436.8	31.8	440	12	BM128059	33
33	431.6	31.4	538	10	BF821434	34
34	404	29.4	436	28	AQ890280	35
35	399.4	29.1	785	12	BI762908	36
36	399.2	29.1	665	13	BY742288	37
37	399	29.1	892	14	CB204861	38
38	394.6	28.7	543	29	CG565104	39
39	394	28.7	413	9	AI422796	40
40	389.2	28.3	698	12	BI906850	41
41	381	27.7	1064	14	CF994566	42
42	379.6	27.6	692	13	BY748962	43
43	374.2	27.3	416	9	AI291866	44
44	372.8	27.2	1319	14	CF594233	45
45	372.6	27.1	584	10	AW917574	

ALIGNMENTS

RESULT 1
BX090012
LOCUS
DEFINITION
BX090012 Soares breast 2NbHst Homo sapiens cDNA clone
IMAGE:154742, mRNA sequence.
ACCESSION
BX090012
VERSION
BX090012.1 GI:27821952
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 776)
AUTHORS
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Pete
Radelof, U., Schneider, D. and Korn, B.
TITLE
Human Unigeneset - RZPD3
JOURNAL
Unpublished (2003)
COMMENT
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGP998E15243.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany

DB	601	GGCRACTAAGAGGGGCTGGACCTTGGCGGCGAGGAAGCAAGAGACTTGGGCGCTAGGCG	
QY	1260	GTTCCCAAATGTGAGGGGGCGAGAAAACAAGCAAGCTCTCTCCCTTGAGAAATTCCTCTC	
DB	661	GTTCCCAAATGTGAGGGGGCGAGAAAACAAGCAAGCTCTCTCCCTTGAGAAATTCCTCTC	
QY	1320	TTTTTAAACAGATATTATTTTATTATTATTATTTGTGACAAATGTGTATAAATGG 1	
DB	721	TTTTTAAACAGATATTATTTTATTATTATTATTATTTGNCACAAATGTTGATAAATGG 1	
RESULT 2			
LOCUS	BI766766	834 bp mRNA linear EST 25	
DEFINITION	60305686GF1 NIH_MGC_122 Homo sapiens cdna clone IMAGE:5206		
ACCESSION	BI766766	mRNA sequence.	
VERSION	BI766766.1	GI:15758344	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 834)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ ; National Institutes of Health, Mammalian Gene Collection Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cdna Library Preparation: Life Technologies, Inc. cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information c found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11517 row: c column: 18 High quality sequence stop: 772.		
FEATURES			
SOURCE			
	1..834	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5206217"		
	/lab_host="DH10B"		
	/clone_lib="NIH_MGC_122"		
	/note="Organ: pooled lung and spleen; Vector: pC		
	Site_1: NotI; Site_2: EcoRV (destroyed); RNA sou		
	anonymous pool of 24 week female lung, 16 week f		
	spleen, and 20-22 week male spleens. Library is		
	primed and directionally cloned (EcoRV site is d		
	upon cloning). Average insert size 1.4 kb, inse		
	range 1-3 kb. Library is normalized and enriched		
	full-length clones and was constructed by C. Gr		
	(Invitrogen). Research genetics tracking code (
	this is a NIH_MGC Library."		
ORIGIN			
Query Match	55.8%	Score 765.8; DB 12; Length 834;	
Best Local Similarity	98.0%	Pred. No. 7,3e-134;	
Matches 818; Conservative	0; Mismatches 12; Indels 5; C		
QY	272	CCGAGGAGGAGCTGGTGGCAGAGAGGAGCACGAGACCCGTCGGAATCGAATCCCCAC	
DB	1	CCCAGGAGGAGCTGGTGGCAGAGGAGGACGAGACCCGTCGGAATCGAATCCCCAC	
QY	332	AAGAAAGCCAGGATCTGCGCCCTTCTCGAAACGAGCTAGTTCGGCCTCGCAGAGT	
DB	61	AAGAAAGCCAGGATCTGCGCCCTTCTCGAAACGAGCTAGTTCGGCCTCGCAGAGT	
QY	392	CTAAGGCCGGGAAAAACACGGGCTTCGAAGAGCGCATCGAGCCCATTTATGAAGTTTCA	

AGCCCGGAAACACGGGCTCGAAGAGCGATCGAGCCCATATGAAGTTTCATCCAC 180
 -TGGACAGGACGGAGCGCAGGAGTGTGGACGGACAGTGTGGCTGGAGGAA 510
 JTGGACAGGACGGAGCGCAGGAGTGTGGACGGACAGTGTGGCTGGAGGAA 240
 AATCAACAGCTCCAGCCCTCTCGCTACAAACCGCAGATCGGGAGTTTATAGTC 570
 AATCAACAGCTCCAGCCCTCTCGCTACAAACCGCAGATCGGGAGTTTATAGTC 300
 AGCTGGCTCTACTACTGTACTGTGACGTGACATTTGATGAGGGAAGCTGTC 630
 AGCTGGCTCTACTACTGTACTGTGACGTGACATTTGATGAGGGAAGCTGTC 360
 GAAGCTGACCTTGTGCTGGATGTGTGCTGGCCCTGGCTGCTGGAGGAATTC 690
 GAAGCTGACCTTGTGCTGGATGTGTGCTGGCCCTGGCTGCTGGAGGAATTC 420
 XACTGGGCGAGTTCCCTCGGGCCCCAGTCTCGCTCTGCGCAGGTGTCTGGGCTG 750
 XACTGGGCGAGTTCCCTCGGGCCCCAGTCTCGCTCTGCGCAGGTGTCTGGGCTG 480
 XCTGGGCGAGGTCCTCCCTCGGGCCCCAGTCTCGCTCTGCGCAGGTGTCTGGGCTG 810
 XCTGGGCGCA-GGTCTCCCTCGGGATCGCACCTCCCTCGGGCCCCATCTCAAG 539
 XCCCTCTCTCACTTCTCGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCC 870
 XCCCTCTCTCACTTCTCGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCC 599
 GTCTGCCAGGTCGGGCTCCCTCGACAGTCTCTGCGGCAACCGGTCCTCC-TC 929
 GTCTGCCAGGTCGGGCTCCCTCGACAGTCTCTGCGGCAACCGGTCCTCCGGG 659
 XACCTCAGCGCTCTTGTCTCCAGACCTGCGCTCCCTCTGAGGGCTGCTGG 989
 XACCTCAGCGCTCTTGTCTCCAGACCTGCGCTCCCTCTGAGGGCTGCTGGG 719
 TCAGTGTCTTCCATCCACATAATACAGTATTCCTTATCTTACAACT- 1048
 TCAGTGTCTTCCATCCACATAATACAGTATTCCTTATCTTACAAAT 779
 CACGCCCCACTCTCACTACTAGTCCCAATCCCTGACCTTTGAGG 1102
 AAACGCCCACTCTCCACTTATAGTCCCAATCCCTGACCTTTGGAGG 834

5 RT_8353983 NIH_MGC_113 Homo sapiens cdna clone IMAGE:6278608
 A sequence.
 5 5.1 GI:21846084
 piens (human)
 piens
 ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 es 1 to 948)
 http://mgi.nci.nih.gov/
 1 Institutes of Health, Mammalian Gene Collection (MGC)
 shed (1999)
 : Robert Strausberg, Ph.D.
 cgabbs-r@mail.nih.gov
 Procurement: Dr. Mark Watson
 Library Preparation: Rubin Laboratory
 library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 quencing by: Agencourt Bioscience Corporation
 distribution: MGC clone distribution information can be
 hrough the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov
 LLCM2466 row: n column: 17

High quality sequence start: 24
 High quality sequence stop: 550.
 Location/Qualifiers
 1..948
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6278608"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_113"
 /notes="Organ: spleen; Vector: pOTB7; Site: 1: XhoI
 EcoRI; cDNA made by oligo-dT priming. Directional
 into EcoRI/XhoI sites using the following 5' adap
 GCACAGAG(G) Library constructed by Ling Hong in
 laboratory of Gerald M. Rubin (University of Cal
 Berkeley) using ZAP-cDNA synthesis kit (Stratagen
 Superscript II RT (Life Technologies). Note: this
 NIH_MGC Library."
 ORIGIN
 Query Match 54.9%; Score 753.2; DB 13; Length 948;
 Best Local Similarity 95.8%; Pred. No. 1.8e-131;
 Matches 807; Conservative 0; Mismatches 30; Indels 7; G
 QY 305 ACCGTCGGAACCTGAATCCCGACAGAGAAAGCCAGGATCCTGCGCCTTCTCTG
 DB 1 ACCGTCGGAACCTGAATCCCGACAGAGAGAGAGAGGATCCTGCGCCTTCTCTG
 QY 365 GACTAGTTCTGGCTCTCGAAGTGCACCTTAAAGCCGGAACACACGGGTCTGAAGA
 DB 61 GACTAGTTCTGGCTCTCGAAGTGCACCTTAAAGCCGGAACACACGGGTCTGAAGA
 QY 425 TCGACGCCCATTTATGAAGTTTCATCCAGACCTGGACAGGACCGAGCGCAGGAGT
 DB 121 TCGACGCCCATTTATGAAGTTTCATCCAGACCTGGACAGGACCGAGC- - -GCAGGT
 QY 485 ACGGACAGTGTGCTGGGAGGAGCCAGAAATCAACAGCTCCAGCCCTCTGCGC
 DB 177 ACGGACAGTGTGCTGGGAGGAGCCAGAAATCAACAGCTCCAGCCCTCTGCGC
 QY 545 ACGGACAGTGTGCTGGGAGTATAGTCAACCGGCTGGGCTCTACTACTGTACTGT
 DB 237 ACGGACAGTGTGCTGGGAGTATAGTCAACCGGCTGGGCTCTACTACTGTACTGT
 QY 605 TGCACCTTTCATGACGAGGAGGCTGTCTACTCAAGCTGGACCTGTGCTGGGATGT
 DB 297 TGCACCTTTCATGAGGAGGAGGCTGTCTACTCAAGCTGGACCTGTGCTGGGATGT
 QY 665 TGGCCCTCTGCTGCTGGGAGGAAATCTCAGCCACTGCGGCGCAGTTCCCTCGGGCC
 DB 357 TGGCCCTCTGCTGCTGGGAGGAAATCTCAGCCACTGCGGCGCAGTTCCCTCGGGCC
 QY 725 TCGGCTCTGCGCAGGTGTCTGGGCTGTGGCCCTGCGGCGCAGGCTCTCCCTGCGG
 DB 417 TCGGCTCTGCGCAGGTGTCTGGGCTGTGGCCCTGCGGCGCAGGCTCTCCCTGCGG
 QY 785 GCACCTCTCCCTGGGCGCATCTCAAGGCTGCCCCCTCTCTCACTACTCTGCGACT
 DB 477 GCACCTCTCCCTGGGCGCATCTCAAGGCTGCCCCCTCTCTCACTACTCTGCGACT
 QY 845 AGGTTCACTGAGGGGCCCTGTGTCTCCCAACAGTCTGCGGAGGCTGCGGCTCCCT
 DB 537 AGGTTCACTGAGGGGCCCTGTGTCTCCCAACAGTCTGCGGAGGCTGCGGCTCCCT
 QY 905 AGCTCTCTGGGCAACCGGTCCTCTGCGCCCAACCTCAGCCGCTCTTTGCTCCAGA-
 DB 597 AGCTCTCTGGGCAACCGGTCCTCTGCGCCCAACCTCAGCCGCTCTTTGCTCCAGA
 QY 964 GCCCTCTCTAGAGGCTGCTGGGCTGTTCAGTGTGTTTCCATCCACATAAA
 DB 657 GCCCTCTCTANAGGCTGCTGGGCTGTTCAGTGTGTTTCCATCCACATAAA
 QY 1024 GTATTCCCACTCTTATCTTACAACTCCGCCCAACCGGCCCACTCTCCCACTCAGTCTC

```

|||||
ATCCCACTCTTATTATTAACAAATCCCCACGCAATTTCTCCACCTCATAGTTCCCC 776
CCCTGACCC--TTTGAGGCCCCAGTGATCTCGATCTCCCTCCCTGCGCAGACAGCCCC 1141
|||||
CCCTGACCCCTTTGAGGGCCCCCAGGGGATCTCAACCTCCCTCCCTGGGCCCAAGACCC 836
|||||
xg 1145
xg 840

259 963 bp mRNA linear EST 15-JUL-2002
XOVRT 8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
RNA sequence.
259
259.1 GI:21782093
sapiens (human)
ycta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 963)
GC http://mgc.nci.nih.gov/.
nal Institutes of Health, Mammalian Gene Collection (MGC)
lished (1999)
ct: Robert Strausberg, Ph.D.
: cgapbs-r@mail.nih.gov
e Procurement: ATCC
Library Preparation: Rubin Laboratory
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Agencourt Bioscience Corporation
e distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
//image.llnl.gov
: LNCM2456 row: 1 column: 13
quality sequence stop: 565.
Location/Qualifiers
1. 963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6274716"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/notes="Organ: salivary gland; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dr priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

54.4%; Score 747.2; DB 13; Length 963;
larity 91.9%; Pred. No. 2.4e-130;
Conservative 0; Mismatches 63; Indels 13; Gaps 7;

TCGGCTCGCAGAGTGCACCTTAAGGCCGGAACACCGGCTCGAAGCGATCGC 428
TCGGCTCGCAGAGTGCACCTTAAGGCCGGAACACCGGCTCGAAGCGATCGC 59
CATTATGAGTTTCATCCAGCAGCTGACAGCAGGAGCGAGCGAGGTGTGACGG 488
|||||
CATTATGAGTTTCATCCAGCAGCTGACAGCAGGAGCGAGCGAGGTGTGACGG 119
AGTGAAGTGGCTGGGAGGAGCCAGATCAACAGCTCCAGCCCTCTCGGTACACCG 548
AGTGAAGTGGCTGGGAGGAGCCAGATCAACAGCTCCAGCCCTCTCGGTACACCG 179

```

```

QY 549 CCAGATCGGGAGTTTATAGTCAACCCGGGTGGGCTCTACTACCTGTACTGTCTCA
Db 180 CCAGATCGGGAGTTTATAGTCAACCCGGGTGGGCTCTACTACCTGTACTGTCTCA
QY 609 CTTTGAATGAGGGAGGCTGTCTACCTGAAAGCTGGAGTCTGCTGGTGGATGGTGY
Db 240 CTTTGAATGAGGGAGGCTGTCTACCTGAAAGCTGGAGTCTGCTGGTGGATGGTGY
QY 669 CTTGCGCTGCTCGAGGAATTTCTCAGCCACTGCGGCCAGTTCCTCGGGCCCCCAK
Db 300 CTTGCGCTGCTCGAGGAATTTCTCAGCCACTGCGGCCAGTTCCTCGGGCCCCCAK
QY 729 CTTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCCAGGGTCTCTCTCGGGATC
Db 360 CTTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCCAGGGTCTCTCTCGGGATC
QY 789 CTTCCCTCGGGCCCATCTCAAGGCTGCCCTCTTCCCTCACCTACTTCGGACTCTTC
Db 420 CTTCCCTCGGGCCCATCTCAAGGCTGCCCTCTTCCCTCACCTACTTCGGACTCTTC
QY 849 TCACTGAGGGCCCTGTGTCTCTCCACAGTGTCTCCAGGTGCGGGCTCCCTCGF
Db 480 TCACTGAGGGCCCTGTGTCTCTCCACAGTGTCTCCAGGTGCGGGCTCCCTCGF
QY 909 CTTGCGGACCCCGTCTCCCTCTGCGGCCACCTCAGCGGCTCTTTGTCTCAGACCT
Db 540 CTTGCGGACCCCGTCTCCCTCTGCGGCCACCTCAGCGGCTCTTTGTCTCAGACCT
QY 969 TCCCTCTAGAGGCTGCTGGGCTGTTCAGCTGTTCATCTCCATCCACATAATACA
Db 600 TCCCTCTAGAGGCTGCTGGGCTGTTCAGCTGTTCATCTCCATCCATTAATACA
QY 1029 CCACCTCTTATCTTCACTCCCTCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCC
Db 660 CCACCTCTTATCTTCACTCCCTCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCC
QY 1088 CC-----TGACCCCTTTGAGGGCCCCCAGTGTCTGAGTCTCCCTCCCTCCCTCC
Db 720 CCCTGACCCCTTTGAGGGCCCCCAGGATCTCGACTCCCTCCCTCCCTCCCTCCCTCC
QY 1144 GGCATT--GTGTTCACTGTACTCTGTGGCAAGG--ATGGGTCCAGAGAGCCCCACT
Db 780 GGCATTGGGTTTCACTGAACTCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
QY 1202 CACTAAGAGGGGCT--GGACCTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
Db 840 ACTTAAAGGGGCTTGGACTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
QY 1257 GGAGTTCCTCAATGTGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
Db 900 GAAATTCCTCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
GAGTTCCTCAATGTGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG

RESULT 5
LOCUS BI819200
DEFINITION 603034614F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:517;
mRNA sequence.
ACCESSION BI819200
VERSION BI819200.1 GI:15930750
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 777)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.

```

613	Db	TGCCGGTCCCTCGACAGCTCTCTGGACACCGGTCCCTCTGCCCCACCCCTCAG
948	QY	TCTTTGTCCAGACTGCCCTCCCTCTAGAGCTCCTGGGCCGTGTTACAGTGT
673	Db	TCTTTGTCCAGACTGCCCTCCCTCTAGAGCTCCTGGGCCGTGTTACAGTGT
1008	QY	ATCCACATAAATACAGTATCCCACTCTTATCTTACAACTCCCC 1052
733	Db	ATCCACAT -AATACAGTATCCCACTCTTATCTTCAATCCCCC 776

RESULT 6	
BM921213	
LOCUS	
DEFINITION	1071 bp mRNA linear EST 12 AGSCNOUT 6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAG 5', mRNA sequence.
VERSION	BM921213
KEYWORDS	BM921213.1 GI:19371592 EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1071)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	Email: cgapbs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information ca found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12786 row: p column: 02 High quality sequence stop: 656.

```

1. 1071
  location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5752561"
  /lab_host="DH10B"
  /clone_lib="NIH MGC 115"
  /note="organ: Pooled brain, lung, testis; Vector:
  pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroy
  source anonymous pool of 6 male brains, age range
  male lung, age 27; and 1 male testis, age 63. Li
  oligo-dT primed and directionally cloned (EcoRV s
  destroyed upon cloning). Average insert size 1.
  insert size range 1-3 kb. Library is normalized s
  enriched for full-length clones and was construc
  Gruber (Invitrogen). Research Genetics tracking
  021. Note: this is a NIH MGC Library."

```

Query Match	52.8%;	Score 725.4;	DB 12;	Length 1071;
Best Local Similarity	93.4%;	Prod. No. 3.1e-126;		
Matches 836;	Conservative 0;	Mismatches 46;	Indels 13;	Gaps 13;
412	GCTCGAAGAGCGCATCGACGCCATTATCAAGTTTCATCCAGCACTCGACACGAGCGG			
dbb	12	GCTCGAAGAGCGCATCGACGCCATTATGAGTTTCATCCAGCACTCGACACGAGCGG		
472	CAGGCGAGGTGTGGA CGGGGACAGTGAGTGGCTGGGAGGAAGCCAGATCAACAGCTCT			
dbb	72	CAGGCGAGGTGTGGA CGGGGACAGTGAGTGGCTGGGAGGAAGCCAGATCAACAGCTCT		
4532	CTCTCTGGCTTCAACACCGCCAGATCGGGAGTTTATAGTTCACCGGGGCTGGGCTCTAT			

TGCGTACACCGCCAGATCGGGAGTTTATAGTACACCGGGCTGGCTCTACTAC 191
 ACTGTCAGGTGCACCTTTCATGAGGGAGAGCTGTCTACTGAAGCTGACCTGTG 651
 ACTGTCAGGTGCACCTTTCATGAGGGAGAGCTGTCTACTGAAGCTGACCTGTG 251
 ATGGTGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 711
 ATGGTGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 311
 AGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 771
 AGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 371
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 831
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 431
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 891
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 491
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 951
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 551
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1011
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 611
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1071
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 671
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1130
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 731
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1186
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 791
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1239
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 851
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1293
 TGGGATCGGACCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 906

11 731 bp mRNA linear EST 11-OCT-2001
 825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5',
 sequence.

11 11.1 GI:16045386

sapiens (human)

sapiens

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ies 1 to 731)

IC http://mgi.nci.nih.gov/

al Institutes of Health, Mammalian Gene Collection (MGC)

ished (1999)

ic: Robert Strausberg, Ph.D.

cgapbs-remail.nih.gov

Procurement: ATCC

Library Preparation: Life Technologies, Inc.

Library Arrayed by: The I.M.G.E. Consortium (LLNL)

Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information c:
 found through the I.M.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM12034 row: b column: 07
 High quality sequence stop: 728.

FEATURES

Location/Qualifiers
 1..731
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5405478"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 90"
 /note="Torgan; liver; Vector: pCMV-SPORT6; Site 1:
 Site 2: SalI; Cloned unidirectionally; oligo-dT 1
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technol
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 51.1%; Score 701.4; DB 12; Length 731;
 Best Local Similarity 99.3%; Pred. No. 9.4e-122;
 Matches 726; Conservative 0; Mismatches 1; Indels 4;
 QY 192 CTTGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1
 Db 1 CTTGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 252
 QY 252 GCTGTCCGCC---CAGGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 61
 Db 61 GCTGTCCGCC---CAGGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 309
 QY 309 GTGGAGTGAATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121
 Db 121 GTGGAGTGAATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
 QY 369 AGTTGGCGCTGGCGAGAGTGCACCTAAAGGCGGAGAGAGAGAGAGAGAGAG 181
 Db 181 AGTTGGCGCTGGCGAGAGTGCACCTAAAGGCGGAGAGAGAGAGAGAGAGAG 429
 QY 429 AGCCCATTTATGAAGTTTATTCACGACCTGGAGACAGAGAGAGAGAGAGAG 241
 Db 241 AGCCCATTTATGAAGTTTATTCACGACCTGGAGACAGAGAGAGAGAGAGAG 489
 QY 489 GACAGTGAAGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
 Db 301 GACAGTGAAGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549
 QY 549 CCAGATCGGGAGTTTATAGTACCCCGGGCTGGCTGTACTACCTGTACTGTCAK 361
 Db 361 CCAGATCGGGAGTTTATAGTACCCCGGGCTGGCTGTACTACCTGTACTGTCAK 609
 QY 609 CTTTGTAGAGAGAGAGTGTCTACTGAGCTGGAGCTGGAGCTGGAGTGGTGG 421
 Db 421 CTTTGTAGAGAGAGAGTGTCTACTGAGCTGGAGCTGGAGCTGGAGTGGTGG 669
 QY 669 CTTGCTGCTGGAGAGATTTCTCAGCACCTGGCGCCAGTTCCCTCGGGCCCCCAG 481
 Db 481 CTTGCTGCTGGAGAGATTTCTCAGCACCTGGCGCCAGTTCCCTCGGGCCCCCAG 729
 QY 729 CTTTGTAGAGTGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 541
 Db 541 CTTTGTAGAGTGTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 789
 QY 789 CTTGCTGCTGGCGCCATCTCAGAGCTGGCGCCCTCTCTACCTACTTGGACTCTTC 600
 Db 600 CTTGCTGCTGGCGCCATCTCAGAGCTGGCGCCCTCTCTACCTACTTGGACTCTTC 849
 QY 849 TCACTGAGGGGGCCCTGTGTCTCCCGACAGTGTCTCCAGGCTGCCGGCTCCCTCGAC 660
 Db 660 TCACTGAGGGGGCCCTGTGTCTCCCGACAGTGTCTCCAGGCTGCCGGCTCCCTCGAC


```

1  TGAACCGACTAGTTCCGCTCGCAGAAAGTGACCTAAAGCGCGGAAACACCGGC 413
2  |||||
3  TGAACCGACTAGTTCCGCTCGCAGAAAGTGACCTAAAGCGCGGAAACACCGGC 360
4  |||||
5  AGAGCGATCCAGCCCATTAAGTTATCCAGACCTCGACAGGAGCGGCA 473
6  |||||
7  AGAGCGATCCAGCCCATTAAGTTATCCAGACCTCGACAGGAGCGGCA 420
8  |||||
9  TGTGTGACGGGACAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGCCC 533
10 TGTGTGACGGGACAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGCCC 480
11 |||||
12 TGTATACCGCCAGATCGGGAGTTATAGTACACCGGGCTGGCTTACTACCT 593
13 TGTATACCGCCAGATCGGGAGTTATAGTACACCGGGCTGGCTTACTACCT 540
14 |||||
15 TGTAGGTGACCTTTGATAGGGGAGGCTGTCTACTGAAGCTGGACTTGTGGT 653
16 TGTAGGTGACCTTTGATAGGGGAGGCTGTCTACTGAAGCTGGACTTGTGGT 600
17 |||||
18 TGTGTGCTGGCTCGCTGCTGGAGGAAATTCAGCCACTGCGGCGCACTTCCCT 713
19 TGTGTGCTGGCTCGCTGCTGGAGGAGGCTGTCTACTGAAGCTGGACTTGTGGT 659
20 |||||
21 TCCAGCTCGGCTTCTGACAGGTGTCTGGGCTGT 751
22 TCCAGCTCGGCTTCTGACAGGTGTCTGGGCTGT 697

```

19 170-avw-m-22-0-UI.r1 NIH_MGC_214 Homo sapiens cdna clone
19 10559117 5', mRNA sequence.

19.1 GI:33203878

piens (human)
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 666)
, M.F., Lennon, G. and Soares, M.B.
zation and subtraction: two approaches to facilitate gene
ry
Res. 6 (9), 791-806 (1996)

: Soares, MB
ated Laboratory for Computational Genomics
ity of Iowa
ton Road, 4156 MEBRF, Iowa City, IA 52242, USA
9 335 8250
9 335 9565
bento-soares@uiowa.edu
Procurement: Mary Hendrix
library preparation: Dr. M. Bento Soares, University of Iowa
quencing by: Dr. M. Bento Soares, University of Iowa
Distribution: Distribution information can be found at
genome.uiowa.edu/distribution/humanfl.html
mer: pYX-5.
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30559117"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_214"
/note="Organ: Lung; Vector: pYX-Asc; Site: 1: EcoR I;
Site_2: Not I; The library was constructed according

mer: pYX-5.
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30559117"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_214"
/note="Organ: Lung; Vector: pYX-Asc; Site: 1: EcoR I;
Site_2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:79
1996. Denatured RNA was size fractionated on a 1%
gel. First strand cDNA synthesis was primed with
primer containing a Not I site. Double strand cDN
size selected according to mRNA size fraction, li
with EcoR I adaptor, digested with Not I and then
directionally into pYX-Asc vector. The library ta
sequence located between the Not I site and the p
is GATAAGGCCA. Tissue was provided by Mary Hendri

ORIGIN

Query Match 47.8%; Score 656.6; DB 14; Length 666;
Best Local Similarity 98.8%; Pred. No. 2.5e-113;
Matches 659; Conservative 0; Mismatches 7; Indels 0; G

```

QY 292 GAGGAGGACGAGACCCGTCGGAACTGAATCCCGACAGCAAGAAAGCCAGGATCC
Db 1 GAGGAGGACCCGACCCGTCGGAACTGAATCCCGACAGCAAGAAAGCCAGGATCC
QY 352 CTTTCTCTGAACCGACTAGTTCCGCTCGCAGAAAGTGCACCTAAAGCCGGAAGC
Db 61 CTTTCTCTGAACCGACTAGTTCCGCTCGCAGAAAGTGCACCTAAAGCCGGAAGC
QY 412 GCTCGAAGAGCGATCGAGCCCATTTATGAAGTTTCATCCAGACCTGGACAGGCGG
Db 121 GCTCGAAGAGCGATCGAGCCCATTTATGAAGTTTCATCCAGACCTGGACAGGCGG
QY 472 CAGGCAAGGTGTGACCGGACAGTGAAGTGGCTGGAGGAGCCAGAAATCAACAGCTC
Db 181 CAGGCAAGGTGTGACCGGACAGTGAAGTGGCTGGAGGAGCCAGAAATCAACAGCTC
QY 532 CTTCTCGGCTACAAACCGCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTA
Db 241 CTTCTCGGCTACAAACCGCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTA
QY 592 CTGTACTGTCAAGTGCACCTTTGATAGGGGAGGCTGTCTACCTGAAGCTGGACTT
Db 301 CTGTACTGTCAAGTGCACCTTTGATAGGGGAGGCTGTCTACCTGAAGCTGGACTT
QY 652 GTGGATGGTGTCTGGCCCTCGCTCGCTCGCTGGAGGAAATTCAGCCACTGGCGCCAG
Db 361 GTGGATGGTGTCTGGCCCTCGCTCGCTCGCTGGAGGAAATTCAGCCACTGGCGCCAG
QY 712 CTCGGGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTGTGGCCCTGGGCGCCAGG
Db 421 CTCGGGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTGTGGCCCTGGGCGCCAGG
QY 772 TCCCTCGGATCGGACACCTCCCTCGCTGGGCCCCATCTCAAGGCTGCCCCCTTCCTCAC
Db 481 TCCCTCGGATCGGACACCTCCCTCGCTGGGCCCCATCTCAAGGCTGCCCCCTTCCTCAC
QY 832 TTCCGACTCTTCCAGGTTTCACTGAGGGGCCCCCTGGTCTCCCGACAGTCGTCGAGGC
Db 541 TTCCGACTCTTCCAGGTTTCACTGAGGGGCCCCCTGGTCTCCCGACAGTCGTCGAGGC
QY 892 GGCTCCCCCTCGACAGCTCTCTGGGACACCGGCTCCCTCTGTGCCCCACCCCTCAGCCGCT
Db 601 GGCTCCCCCTCGACAGCTCTCTGGGACACCGGCTCCCTCTGTGCCCCACCCCTCAGCCGCT
QY 952 TGCTCC 957
Db 661 TGCTCC 666

```

RESULT 12
BI966060/c

LOCUS
DEFINITION

BI966060 609 bp mRNA linear EST 12-
ie72904.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sa
cdna clone IMAGE:5672623 3' similar to TR:054907 054907 TNE
WEAK INDUCER OF APOPTOSIS ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

BI966060
BI966060.1 GI:16340465

EST.


```

|||||
TCTGGGACCGCGTCCCTCTGYYCCACCTCAGCGCTCTTGTCTCAGACCTG 180
|||||
CCCTCTAGAGCTGCTGGGCTGTTCACGTTTTCATCCACATAAATACAG 1024
|||||
CCCTCTAGAGCTGCTGGGCTGTTCACGTTTTCATCCACATAAATACAG 240
|||||
CCACTCTATCTTCAACTCTCCCAACCGCCCACTCTCCACCTCACTAGTCCCA 1084
|||||
CCACTCTATCTTCAANNYYCCACCGCCCACTCTCCACCTCACTAGTCCCA 300
|||||
TGACCTTTGAGGCCCCAGTGTCTGACTTCCCTCCCTGGCCACAGACCCCAAG 1144
|||||
TGACCTTTGAGGCCCCAGTGTCTGACTTCCCTCCCTGGCCACAGACCCCAAG 360
|||||
TGCTTCACTGTACTGTGGGCAAGGATGGTCCAGAGACCCCACTTCAAGCAC 1204
|||||
GTGTTCACCTGTACTGTGGGCAAGGATGGTCCAGAGACCCCACTTCAAGCAC 420
|||||
GGGCTGACCTGTGGGCAAGGATGGTCCAGAGACCCCACTTCAAGCAC 1264
|||||
GGGCTGACCTGTGGGCAAGGATGGTCCAGAGACCCCACTTCAAGCAC 480
|||||
GTGAGGGCGGAGAAACAAGCAAGCTCTCTCTTGAATTTCCCTGTGATTTT 1324
|||||
GTGAGGGCGGAGAAACAAGCAAGCTCTCTCTTGAATTTCCCTGTGATTTT 540
|||||
AGATATTATTTTATTATTATTGTGACAAATGTTGATAATGG 1373
|||||
AGATATTATTTTATTATTATTGTGACAAATGTTGATAATGG 589

```

```

9 567 bp mRNA linear EST 06-NOV-2002
.y1 Human Retinal pigment epithelium/choroid cDNA
malized, unamplified): cs Homo sapiens cDNA clone cs80h07
9
9.1 GI:24733297

```

```

piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 567)
n,J.W., Bouffard,G., Smith,D. and Peterson,K.
ed sequence tag analysis of human RPE/choroid for the
Project: Over 6000 non-redundant transcripts, novel genes
ice variants
s. 8 (4), 205-220 (2002)
0
0

```

```

: Wistow G
on Molecular Structure and Function
1 Eye Institute
NIH, Bethesda, MD 20892-2740, USA
1 402 3452
1 496 0078
graeme@helix.nih.gov
80 row: h column: 07
mer: M13RPI reverse primer (ABI).
Location/Qualifiers
1. .567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cs80h07"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retinal pigment epithelium/choroid cDNA

```

```

(Un-normalized, unamplified): cs"
/notes="Organ: Eye; Vector: pCMVSPORT6; Two differ
eyes (75-80 years old) yielded approximately 600 t
dissected RPE/choroid tissue. This in turn yield
of total RNA and 7 ug of mRNA. A directionally cl
library in the pCMVSPORT6 vector was constructed i
Technologies (Rockville, MD; now part of Invitroge
essentially following the protocols of the SuperSk
Plasmid System (Invitrogen Corp.
<http://www.invitrogen.com/>). The library code
designation was cs. For this library, cDNA insert:
cloned into the NotI/MluI sites of the vector. ES:
analysis was performed on the unamplified library
NIH Intramural Sequencing Center (NISC)."
```

ORIGIN

```

Query Match 41.3%; Score 567; DB 14; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.8e-96;
Matches 567; Conservative 0; Mismatches 0; Indels 0; G:
QY 55 CTCGGGTCCCGGATGGGGGGGGGTGAGGACGACACGCCCCCGCCCATGGC
Db 1 CTCGGGTCCCGGATGGGGGGGGGTGAGGACGACACGCCCCCGCCCATGGC
QY 115 CGTCGAGCCAGAGCGAGGGGGGGCGCGGGGGGAGCGCGCCCTGTGTGGT
Db 61 CGTCGAGCCAGAGCGGAGGGGGGGCGCGGGGGGAGCGCGCCCTGTGTGGT
QY 175 CTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG
Db 121 CTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG
QY 235 TTGGGGAGCCGGGATCGCTGTCGCCCGGAGGAGCTGCCCGAGGAGCTGGTGGC
Db 181 TTGGGGAGCCGGGATCGCTGTCGCCCGGAGGAGCTGCCCGAGGAGCTGGTGGC
QY 295 GAGGACCCAGACCCCGTCGGAATGAATCCCGAGACAGAAAGCCAGGATCCTGC
Db 241 GAGGACCCAGACCCCGTCGGAATGAATCCCGAGACAGAAAGCCAGGATCCTGC
QY 355 TTCCTGAACCGACTAGTTCCGGCTCGAGAGTGACCTAAAGGCCGGAACACCG
Db 301 TTCCTGAACCGACTAGTTCCGGCTCGAGAGTGACCTAAAGGCCGGAACACCG
QY 415 CGAAGAGCGATCGAGCCCATTTATGAAGTTTATCCAGACCTGGACAGCCGAGC
Db 361 CGAAGAGCGATCGAGCCCATTTATGAAGTTTATCCAGACCTGGACAGCCGAGC
QY 475 GCAGGTGTGAGCGGACAGTGTAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAG
Db 421 GCAGGTGTGAGCGGACAGTGTAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAG
QY 535 CTGGCTACAACCGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTTACTAC
Db 481 CTGGCTACAACCGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTTACTAC
QY 595 TACTGTGAGTGCACTTTGATGAGGG 621
Db 541 TACTGTGAGTGCACTTTGATGAGGG 567

```

RESULT 15

```

BO884231
LOCUS
DEFINITION
BO884231
IMAGE:6197488 5', mRNA sequence.
ACCESSION
BO884231
VERSION
BO884231.1 GI:22276239
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

ases 1 to 940)
3C http://mgc.nci.nih.gov/.
nal Institutes of Health, Mammalian Gene Collection (MGC)
lished (1999)
ct: Robert Strausberg, Ph.D.
: cgapsb-f@mail.nih.gov
e Procurement: Dr. James R. Lupski
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Sequencing by: Agencourt Bioscience Corporation
e distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LNL at:
//image.lnl.gov
: LLM13607 row: j column: 17
quality sequence stop: 453.
Location/Qualifiers
1. 940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6197488"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sciatic nerve"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCCAGCCACGCGTCCG-3' and
5'-GACTAGTTCAGATCCGAGCGCCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
40.9%; Score 562; DB 13; Length 940;
larity 96.5%; Pred. No. 1.8e-95;
Conservative 0; Mismatches 15; Indels 9; Gaps 8;
IACCGTCGGAACCTGAATCCCGACAGAGAAAGCCAGGATCCTGGCCTTTCCCTG 360
IACCGTCGGAACCTGAATCCCGACAGAGAAAGCCAGGATCCTGGCCTTTCCCTG 60
IAGCTAGTTCCGGCTCGCAGAGTGCACCTAAAGCGCGGAACACGGGCTCGAAGA 420
IAGCTAGTTCCGGCTCGCAGAGTGCACCTAAAGCGCGGAACACGGGCTCGAAGA 120
ITCGAGCCCATTTATGAAGTTTCATCCAGACCTGGACAGGAGCGAGCGAGGT 480
ITCGAGCCCATTTATGAAGTTTCATCCAGACCTGGACAGGAGCGAGCGAGGT 180
IACGGACACAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGCCCTCTGGCC 540
IACGGACACAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGCCCTCTGGCC 240
IACGGCCAGATCGGGGAGTTTATAGTCACCGGGCTGGGCTCTACTACTGTACTGT 600
IACGGCCAGATCGGGGAGTTTATAGTCACCGGGCTGGGCTCTACTACTGTACTGT 300
ITGCACTTTTATGAGGGGAGGCTGTCTACCTGAAGCTGGACTGTGCTGGTGGATGGT 660
ITGCACTTTTATGAGGGGAGGCTGTCTACCTGAAGCTGGACTGTGCTGGTGGATGGT 360
ITGGCCCTTGGCTGCTCGAGGAAATTCCTAGCCACTGGCGGCGAGTTCCCTCGGGGCC 720
ITGGCCCTTGGCTGCTCGAGGAAATTCCTAGCCACTGGCGGCGAGTTCCCTCGGGGCC 420
ITCCGCTCTG-CCAGGTGTCTGGGCTGTGGCCCTGGCGCCA-GGGTCTCCCTGC 778
ITCCGCTCTGCCCCAGGTGTCTGGGCTGTGGCCCTGGCGGCCAGGGGCCCTCCCTGC 480

```

```

QY 779 GGATCCGCACCCCTCCCTGGGGCCCATCTCAAGGCTG-CCCCCTTCTCCTACCTAC-
Db 481 GGATCCGCACCCCTCCCTGGGGCCCATCTCAAGGCTGCCCCCTTTCTCCTACTACT
QY 837 ACTCTTCCAGGTTTCACTGAGGGGCCCTTGGTCTCTCCCAACAGTCTGT-CCCCAGGCTGC
Db 541 ACTCTTCCAGGTTTCACTGAGGGGCCCTTGGTCTCTCCCAACAGTCTGT-CCCCAGGCTGC
QY 895 TCCCTCTGACAGTCTCT-GGGACACCGGTTCCCTCTGCCCCCACC-TCAGCCG
Db 601 TCCCTCTGACAGTCTCTGAGGAAACCGGTTCCCTCTGCCCCCACCCTCAGGCGG
QY 952 TGCTCCAGACCTGGCCCTCCCTCT 975
Db 661 TGGTCCAGAACTGGCCCTCCCTCT 684

```

Search completed: April 7, 2004, 23:15:25
Job time : 3339.48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

ic search, using sw model

ril 7, 2004, 17:30:20 ; Search time 5268.84 Seconds
(without alignment)
11294.703 Million cell updates/sec

-09-245-198a-3

73

atgcattgttagactttga.....gacaaaatgtgataaatgg 1373

ENTITY_NUC

pop 10.0 , Gapext 1.0

70272 seqs, 21671516995 residues

ts satisfying chosen parameters: 6940544

3th: 0

3th: 2000000000

inimum Match 0%

aximum Match 100%

isting first 45 summaries

enEmbl.*

gb_ba.*

gb_htg.*

gb_in.*

gb_on.*

gb_ov.*

gb_pat.*

gb_ph.*

gb_pl.*

gb_pr.*

gb_ro.*

gb_sts.*

gb_sy.*

gb_un.*

gb_vi.*

em_ba.*

em_fun.*

em_hum.*

em_in.*

em_mu.*

em_or.*

em_ov.*

em_pat.*

em_ph.*

em_pl.*

em_ro.*

em_sts.*

em_un.*

em_vi.*

em_htg_hum.*

em_htg_inv.*

em_htg_other.*

em_htg_mus.*

em_htg_pln.*

em_htg_rnd.*

em_htg_mam.*

em_htg_vrt.*

em_sy.*

em_htgo_hum.*

em_htgo_mus.*

em_htgo_other.*

the number of results predicted by chance to have a

score greater than or equal to the score of the result being pri
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1373	100.0	1373	6	BD062758	BD062758 A
2	1325	96.5	1421	6	BD090952	BD090952 Ap
3	1320.2	96.2	1353	6	AX201324	AX201324 Se
4	1320.2	96.2	1353	9	AY358870	AY358870 Hc
5	1320.2	96.2	1368	9	AF055872	AF055872 Hc
6	1285	93.5	1306	9	AF030099	AF030099 Hc
7	1226.4	89.3	1236	6	ARI40407	ARI40407 Se
8	1226.4	89.3	1236	6	BD057124	BD057124 Me
9	1096.8	79.9	1642	9	BC019047	BC019047 Hc
C 10	768.2	56.0	60268	9	AC016876	AC016876 Hc
11	754	54.9	218485	2	AC127470	AC127470 Pa
12	683.4	49.8	1239	10	AF030100	AF030100 M
13	629.2	45.8	898	6	AX180714	AX180714 Se
14	614.6	44.8	1168	6	BD062757	BD062757 A
15	553	40.3	1816	9	AY081051	AY081051 Hc
16	364.2	26.5	148555	2	AC126921	AC126921 Bc
C 17	309.6	22.5	203083	2	AC069459	AC069459 Mu
C 18	309.6	22.5	234182	10	AL603707	AL603707 M
C 19	302.2	22.0	130254	2	AC136195	AC136195 Ra
C 20	302.2	22.0	165316	2	AC119115	AC119115 Ra
C 21	302.2	22.0	223877	2	AC098923	AC098923 Ra
C 22	302.2	22.0	225077	2	AC136563	AC136563 Ra
C 23	283.4	20.6	149736	2	AC126239	AC126239 Fe
C 24	267	19.4	234801	2	AC118309	AC118309 Ra
C 25	246.4	17.9	180222	2	AC130192	AC130192 Su
C 26	242.4	17.7	176258	2	AC126925	AC126925 Ca
C 27	193.4	14.1	195	6	AX379024	AX379024 Se
C 28	184.2	13.4	212093	2	AC126237	AC126237 Ca
C 29	87.2	6.4	7218	6	I66494	I66494 Sequ
C 30	77.6	5.7	125020	9	AF429315	AF429315 Hc
C 31	74.8	5.4	303091	2	AC084799	AC084799 Mu
C 32	70.8	5.2	73948	2	AC022556	AC022556 Ho
C 33	70.6	5.1	144979	2	AC016280	AC016280 Hc
C 34	70.4	5.1	220469	2	AC074307	AC074307 Mu
C 35	70.2	5.1	298166	2	AC087563	AC087563 Ho
C 36	69.2	5.0	936	8	CNS01AYK	AL113556 Bc
C 37	69.2	5.0	82897	2	AC022595	AC022595 Hc
C 38	68.4	5.0	202083	2	AC023833	AC023833 Mu
C 39	68.2	5.0	110737	2	AC011105	AC011105 Hc
C 40	68.2	5.0	265537	2	AC087228	AC087228 Mu
C 41	68.2	5.0	266935	2	AC144767	AC144767 Mu
C 42	68	5.0	86784	2	AC138362	AC138362 Mu
C 43	68	5.0	300695	2	AC079431	AC079431 Mu
C 44	67.6	4.9	626	8	CNS018RP	AL110716 Bc
C 45	67.2	4.9	3163	6	AX647321	AX647321 Se

ALIGNMENTS

RESULT 1	BD062758	1373 bp	DNA	linear	PAT 27
LOCUS	BD062758				
DEFINITION	A tumor necrosis factor related ligand.				
ACCESSION	BD062758				
VERSION	BD062758.1	GI:22608361			
KEYWORDS	JP 2001505407-A/2.				
SOURCE	unidentified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 1373)				
AUTHORS	Chicheportiche, Y. and Browning, J. L.				
TITLE	A tumor necrosis factor related ligand				
JOURNAL	Patent: JP 2001505407-A 2 24-APR-2001;				
	BIOTEN INC, THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GE				

[illegible]

Location/Qualifiers

```
1..1421
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

          96.5%;   Score 1325;   DB 6;   Length 1421;
          rity 98.5%;   Pred. No. 1.1e-232;
          nservative 0;   Mismatches 20;   Indels 0;   Gaps 0;

ATTTCGCGCCGCGGCTCCCTCCCTCCCGATCCCTCGGTCCTCGGATCGGCGGAGTGGGGGG 76
ATCCCTCAGCTCGACCCACGCGTCCGCGATCCCTCGGTCCTCGGATCGGCGGAGTGGGGGG 62
AGGAGGACACAGCCCCCGCCCGATGCGCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 136
AGGAGGACACAGCCCCCGCCCGATGCGCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 122
AGGAGGACACAGCCCCCGCCCGATGCGCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 196
CGGGGGAGCGCGGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182
GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
CAGGAGCTGCCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 316
CAGGAGCTGCCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302
CCCCAGACAGAAAGACGAGATCCTCGGCTTTCCTGAAACCGAGTCTGCTGCGG 376
CCCCAGACAGAAAGACGAGATCCTCGGCTTTCCTGAAACCGAGTCTGCTGCGG 362
AGAGTGACCTAAAGCGCGGAAACACACGCGCTCGAAGAGGATCGAGCCCAAT 436
AGAGTGACCTAAAGCGCGGAAACACACGCGCTCGAAGAGGATCGAGCCCAAT 422
GTTTCATCCAGGCTGGACAGACGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 496
GTTTCATCCAGGCTGGACAGACGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 482
TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 556
TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
TTTATAGTACCGCGGCTGGGCTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
TTTATAGTACCGCGGCTGGGCTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
AAGGCTGTCTACCTGAAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
AAGGCTGTCTACCTGAAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
AGGAAATTTCTAGGCACTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
AGGAAATTTCTAGGCACTCGCGGAGTTCCTCGGCGGCGGCGGCGGCGGCGGCGGCG 722
TCTGGGCTGTGGGCTCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 796
TCTGGGCTGTGGGCTCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
CATCTCAAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
CATCTCAAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
CTGGTCTCCCACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916
CTGGTCTCCCACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 902
GTCCCTCTGCCCCACCTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
GTCCCTCTGCCCCACCTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 962
```


[illegible]

[illegible]

CTCACTAGTCCCAATCCCTGACCCCTTTGAGGCCCCCACTGATCTCGACTCCCCC 1020
 CCACAGACCCCAAGGAGGATTTGTTCTACTGTCTGTGGCAAGGATGGTCCAG 1185
 CCACAGACCCCAAGGAGGATTTGTTCTACTGTCTGTGGCAAGGATGGTCCAG 1080
 CCCACCTTCAGGCACTAAGAGGGCTGAGCTGTGGCGCAGAGCCAAAGAGACTG 1245
 CCCACCTTCAGGCACTAAGAGGGCTGAGCTGTGGCGCAGAGCCAAAGAGACTG 1140
 TAGGCCAGGAGTTTCCCAATGTGAGGGCGGAGAAACAAGACAGCTCCCTCTGA 1305
 TAGGCCAGGAGTTTCCCAATGTGAGGGCGGAGAAACAAGACAGCTCCCTCTGA 1200
 CCCCTGTGATTTTAAACAGATATATTT 1341
 CCCCTGTGATTTTAAACAGATATATTT 1236

7 1642 bp mRNA linear PRI 04-OCT-2003
 iplens tumor necrosis factor (ligand) superfamily, member 12,
 ipt variant 2, mRNA (cdna clone MGC:20669 IMAGE:4766071),
 e cds.
 7
 7.2 GI:33874709
 piens (human)
 piens
 ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Butheria; Primates; Catarrhini; Hominidae; Homo.
 es 1 to 1642)
 erg.R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 r.R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 l.S.F., Zeeberg,B., Bueow,K.H., Schaefer,C.F., Bhat,N.K.,
 R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 uko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 on M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 T.E., Brange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 i.P., Prongate,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 1.R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 1.K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 C.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 i.G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 eld,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
 A., Schein,J.E., Jones,S.J. and Marra,M.A.
 on and initial analysis of more than 15,000 full-length
 id mouse cDNA sequences
 t.I. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

es 1 to 1642)
 rg,R.
 dmission
 d (07-DEC-2001) National Institutes of Health, Mammalian
 lection (MGC), Cancer Genomics Office, National Cancer
 e, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 Project URL: <http://mgs.nci.nih.gov>
 9, 2003 this sequence version replaced gi:17512138.
 MGC help desk
 gapbs-x@mail.nih.gov
 rocurment: Louis Staudt
 rary Preparation: Rubin Laboratory
 rary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 encing by: National Institutes of Health Intramural
 ng Center (NISC),
 burg, Maryland;

Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgcnhgr@nisc.nih.gov
 Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Broc
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legas
 Maduro,Q.L., Masiello,C., Maekeri,B., Mastrian,S.D., McClos
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touch
 Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggin
 Young,A., Zhang,L.H. and Green,E.D.

Clone distribution: MGC clone distribution information can
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.ll>
 Series: IRAL Plate: 30 Row: p Column: 5
 This clone was selected for full length sequencing because
 passed the following selection criteria: matched mRNA gi:

FEATURES

source

1. .1642
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:20669 IMAGE:4766071"
 /cissue_type="Primary B-Cells from Tonsils"
 /clone_lib="NIH MGC_48"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1. .1642
 /gene="TNFSF12"
 /note="synonyms: TWEAK, DR3LG, APO3L, MGC20669"
 /db_xref="LocusID:8742"
 /db_xref="MIM:602655"
 97. .501
 /codon_start=1
 /product="tumor necrosis factor (ligand) superfam
 member 12, isoform 2"
 /protein_id="AAH19047.1"
 /db_xref="GI:17512139"
 /db_xref="LocusID:8742"
 /translation="MAAFRRSRRRRGEPGTALIVPLALGLGLALCIG
 GSRASLQAQPAQELVAEDQPSSELPQTEESQDPAPFLNRLVRPRS
 ARRATAAHYEVHPRPGDGAQADGYITCLRP"

gene

1. .1642

CDS

Query Match 79.9%; Score 1096.8; DB 9; Length 1642;
 Best Local Similarity 84.8%; Pred. No. 6.9e-191;
 Matches 1347; Conservative 0; Mismatches 2; Indels 239; Ga
 QY 25 TCCGCCCGCGGCTCCCGCCCGGATCCCTCGGCTCCCGGATGGGGGGCGGTG
 Db 16 TCCGCCCGCGGCTCCCGCCCGGATCCCTCGGCTCCCGGATGGGGGGCGGTG
 QY 85 CAGCACAGCCCCCGCCCGGATCCCTCGGCTCCCGGATGGGGGGCGGTG
 Db 76 CAGCACAGCCCCCGCCCGGATCCCTCGGCTCCCGGATGGGGGGCGGTG
 QY 145 GGGGAGCGGGCAACCGCTCTGGTCCCGCTCGGCTCCCGGATGGGGGGCGGTG
 Db 136 GGGGAGCGGGCAACCGCTCTGGTCCCGCTCGGCTCCCGGATGGGGGGCGGTG
 QY 205 TGCTCGGCTCTCGTGGCGGCTGGTTCAGTTTGGGAGCGGGGATCGTGTCCGCC
 Db 196 TGCTCGGCTCTCGTGGCGGCTGGTTCAGTTTGGGAGCGGGGATCGTGTCCGCC
 QY 265 GAGCTCTCCCGAGGAGCTGGTGGCAGAGGAGCCAGGACCGCTCGGAACCTGAATC
 Db 256 GAGCTCTCCCGAGGAGCTGGTGGCAGAGGAGCCAGGACCGCTCGGAACCTGAATC
 QY 325 CAGACAGAAAGAGCAGGATCTCGCTTTCCTGAACCGACTAGTTCGGCTCCCA
 Db 316 CAGACAGAAAGAGCAGGATCTCGCTTTCCTGAACCGACTAGTTCGGCTCCCA
 QY 385 AGTGACCTTAAAGCGCGGAACAACCGGCTCCAGAGCGATCGAGCCCATATGAAG

GCACCTAAAGGCCGCGAAACACGGGCTCGAAGAGCGATCGCAGCCATTATGAAGTT 433
CCACGACTCGACACGAGCGAGCGCAGG----- 475
CCACGACTCGACACGAGCGAGCGCAGGATGAGGTTACACAACCTGTCTGAGG 495
----- 475
TGAGATACTAAGTGGTGGAGCCAAGATTGAAACCCAGCTAGNATGTGCTGTGTACT 555
----- 475
ATGGTCTGCTGATGAACGGCAGGTGGCTGTGTCAGAGGGTGAGGGTCCATGCAGGG 615
----- 475
ATATCAAAAGGGGAGAGGAAGTTTCCAGNAAAGAGAGGACACATCTCCCCACCA 675
-----CAGTGTGGACGGGACAGTGAAGTGGCTGGG 505
CAGAGGGTCAAAGGAGAAACAGAGACCAGGTGTGGACGGGACAGTGAAGTGGCTGGG 735
AGCCAGATCAACAGCTCCAGCCCTCTGGCTACAACCGCCAGATCGGGGAGTTTA 565
AGCCAGATCAACAGCTCCAGCCCTCTGGCTACAACCGCCAGATCGGGGAGTTTA 795
CACCCGGCTGGGCTCTACTACTGTACTGTCAAGTGTCACTTTGATGAGGGGAAGG 625
CACCCGGCTGGGCTCTACTACTGTACTGTCAAGTGTCACTTTGATGAGGGGAAGG 855
CTACTGAAGCTGGACTTGTCTGTGTGATGTGTGTGGCCCTGCGCTGCGCTGAGG 685
CTACTGAAGCTGGACTTGTCTGTGTGATGTGTGTGGCCCTGCGCTGCGCTGAGG 915
CTCAGCCACTCGGGCAGTTCCCTCGGGCCCGAGCTCCGCTCTGCCAGGTGTCTG 745
CTCAGCCACTCGGGCAGTTCCCTCGGGCCCGAGCTCCGCTCTGCCAGGTGTCTG 975
GTTGGCCCTCGGGCAGGGTCTTCCTCGGGATCCGACACCTCCCTGGGGCCCATC 1035
GTTGGCCCTCGGGCAGGGTCTTCCTCGGGATCCGACACCTCCCTGGGGCCCATC 865
GGCTGGCCCTTCCTCACCCTACTTCGGACTCTTCAGAGTTCACTGAGGGCCCTCG 1095
GGCTGGCCCTTCCTCACCCTACTTCGGACTCTTCCTCAGGTTCACCTGAGGGCCCTCG 925
CCACAGTCTGCCAGGCTGCCGCTCCCTTCGACAGCTCTCTGGGCAACCGGCTCC 1155
CCGAGTCTGCCAGGCTGCCGCTCCCTTCGACAGCTCTCTGGGCAACCGGCTCC 985
GGCCCACTCTGACCGCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCTGCC 1215
GGCCCACTCTGACCGCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCTGCC 1045
CTGTTCAGGTGTTTCCATCCCATATAATACAGTATTCCTTATCTTACA 1275
CTGTTCAGGTGTTTCCATCCCATATAATACAGTATTCCTTATCTTACA 1105
CCACCGGCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCC 1335
CCACCGGCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCC 1165
GATCTCGACTCCCTCGGCCACAGACCCCGAGGGCATTTGTTCATCTGTACTCT 1395
GATCTCGACTCCCTCGGCCACAGACCCCGAGGGCATTTGTTCATCTGTACTCT 1225
CAAGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCG 1455
CAAGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCG 1285
AAGCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAATGTGAGGGGCGAGAAAC 1515
AAGCCAAAGAGACTGGGCTAGGCCAGGAGTTCCCAATGTGAGGGGCGAGAAAC

1286 AAGCAAGCTCTCCCTTCAGAAATTCCTGTGGATTTTAAACAGATATTATT
 Db AAGCAAGCTCTCCCTTCAGAAATTCCTGTGGATTTTAAACAGATATTATT
 1516 AAGCAAGCTCTCCCTTCAGAAATTCCTGTGGATTTTAAACAGATATTATT
 Qy ATTTATTGTGACAAATGTTGTAATGG 1373
 Db ATTTATTGTGACAAATGTTGTAATGG 1603

RESULT 10
 AC016876/c
 LOCUS
 DEFINITION Homo sapiens chromosome 17, clone RP11-186B7, complete se
 AC016876
 ACCESSION AC016876.10 GI:24431829
 VERSION HTG.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., An
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
 JOURNAL 1 (bases 1 to 60268)
 REFERENCE
 AUTHORS Homo sapiens chromosome 17, clone RP11-186B7
 2 (bases 1 to 60268)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., An
 Baldwin,J., Barna,N., Becker,T., Boguslavsky,L., Boukhga
 Brown,A., Castle,A., Colangelo,M., Collins,S., Connelan,L.,
 Cooke,P., DeArellano,K., Dewar,K., Domino,M., Doolan,L.,
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearnford,A., H
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., K
 Lehotzky,J., Lieu,C., Locke,K., Macdonald,P., Marguis,N.,
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Seve
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., W
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (08-DEC-1999) Whitehead Institute/MIT Center for
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 60268)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anders
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhga
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dod
 Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norma
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,I
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Ro
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Sevi
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Ta
 Tesfaye,S., Theodore,B., Topham,K., Travers,M., Vassiliev,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zaitn
 Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (08-OCT-2002) Whitehead Institute/MIT Center for
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 60268)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderse
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhga
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dod
 Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norma
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,I
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Ro
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Sevi
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Ta
 Tesfaye,S., Theodore,B., Topham,K., Travers,M., Vassiliev,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zaitn
 Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (08-OCT-2002) Whitehead Institute/MIT Center for
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 60268)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderse
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhga
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dod
 Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norma
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,I
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Ro
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Sevi
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Ta
 Tesfaye,S., Theodore,B., Topham,K., Travers,M., Vassiliev,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zaitn
 Zembek,L., Zimmer,A. and Zody,M.

A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 ad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
 ws, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,
 V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
 on, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
 Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,
 S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 ovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 s, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Submission
 ted (31-OCT-2002) Whitehead Institute/MIT Center for Genome
 ch, 320 Charles Street, Cambridge, MA 02141, USA
 31, 2002 this sequence version replaced gi:23592141.
 peats were identified using RepeatMasker:
 A.F.A. & Green, P. (1996-1997)
 /ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 nter: Whitehead Institute/ MIT Center for Genome Research
 nter code: WIBR
 o site: http://www-seq.wi.mit.edu
 ntract: sequence submissions@genome.wi.mit.edu
 ----- Project Information
 iter project name: L3849
 iter clone name: 186_B_7

ie first 60.3 kilobases of this clone are being submitted.
 inander overlaps accession number AC113189 [WICGR project

Location/Qualifiers

1. .60268
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-186B7"
 /clone_lib="RPCI-11 Human Male BAC"
 complement(7920..8209)
 /rpt_family="AluSg"
 complement(8822..9119)
 /rpt_family="AluSx"
 complement(9356..9663)
 /rpt_family="AluSg"
 complement(12421..12733)
 /rpt_family="AluSx"
 14360..14438
 /rpt_family="(TGG)n"
 complement(17883..18086)
 /rpt_family="MIR"
 complement(18188..18483)
 /rpt_family="AluSg"
 22454..22535
 /rpt_family="(TA)n"
 24279..24530
 /rpt_family="MIR"
 complement(24517..24655)
 /rpt_family="MIR"
 24839..24954
 /rpt_family="L2"
 complement(24973..25268)
 /rpt_family="AluY"
 complement(25273..25768)
 /rpt_family="MER21B"
 complement(25769..26057)
 /rpt_family="AluSg"
 complement(26058..26334)
 /rpt_family="MER21B"
 26344..26838
 /rpt_family="L2"
 complement(26839..27139)

repeat_region
 /rpt_family="AluSg"
 27140..27185
 /rpt_family="L2"
 27405..27430
 /note="<30 qual SNGL region"
 27529..27550
 /note="single clone coverage"
 complement(27552..27689)
 /rpt_family="FLAM_C"
 27568..27572
 /note="<30 qual SNGL region"
 complement(27690..28000)
 /rpt_family="AluSx"
 repeat_region
 complement(28005..28312)
 /rpt_family="AluSp"
 30022..30088
 /rpt_family="L2"
 30817..30846
 /rpt_family="GC_rich"
 repeat_region
 31023..31072
 /rpt_family="(TCC)n"
 32341..32378
 /rpt_family="(CCG)n"
 32571..32601
 /rpt_family="GC_rich"
 repeat_region
 37977..38092
 /rpt_family="L1MB5"
 repeat_region
 38094..38400
 /rpt_family="AluSx"
 38910..39189
 /rpt_family="L2"
 complement(39218..39381)
 /rpt_family="AluSg/x"
 39394..39457
 /rpt_family="L2"
 39458..39748
 /rpt_family="AluY"
 39593..39597
 /note="<30 qual SNGL region"
 39749..39759
 /rpt_family="L2"
 repeat_region
 39760..40068
 /rpt_family="AluY"
 repeat_region
 40069..40152
 /rpt_family="L2"
 repeat_region
 40350..40630
 /rpt_family="L2"
 40718..40905
 /rpt_family="MIR"
 repeat_region
 40926..41025
 /rpt_family="L2"
 repeat_region
 41026..41341
 /rpt_family="AluSp"
 41342..41745
 /rpt_family="L2"
 41746..42046
 /rpt_family="AluSx"
 42047..42288
 /rpt_family="L2"
 42289..42581
 /rpt_family="AluSx"
 repeat_region
 42582..42730
 /rpt_family="L2"
 repeat_region
 43730..43859
 /rpt_family="MER94"
 44018..44132

Query Match

Best Local Similarity 56.0%; Score 768.2; DB 9; Length 60268;
 Matches 770; Conservative 0; Mismatches 3; Indels 0; Gs

Qy

601 CAGGTGCACCTTGTATGAGGAGGCTGTACCTGAAGCTGGACTTGCCTGCGAT

06:25:14 2004

us-09-245-198a-3.oligo.rge

Query Match 1.9%; Score 26; DB 6; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.022; 0; Indels 0;
Matches 26; Conservative 0; Mismatches 0;

095 486 bp DNA linear PAT 17-JUL-2003
ace characteristic to gene transcription controlled by
ia.
095 1.486
095.1 GI:33033865
02525081-A/10.
sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 486)
P., Skalliter, R. and Feinstein, E.
nce characteristic to gene transcription controlled by hypoxia
t: JP 2002525081-A 10 13-AUG-2002;

BIOTECH INC
Homo sapiens (human)
JP 2002525081-A/10
13-AUG-2002
27-AUG-1998 JP 2000571058
27-AUG-1998 US 60/098158, 05-MAY-1999 US 60/132684 PI
INAT, RAMI, SKALLITER, ELENA FEINSTEIN
C12N15/09, A61K31/711, A61K45/00, A61K48/00, A61P9/00, A61P9/10, PC
F43/00, G01N33/53, G01N33/566, G01N33/58, G01N33/68, G01N37/00,
C12Q1/68, G01N33/53, G01N33/566, G01N33/58, G01N33/68, G01N37/00,
C12N15/00
Sequence characteristic to gene transcription controlled by
hypoxia

Key Location/Qualifiers
source 1. .486
/organism="Homo sapiens (human)"
Location/Qualifiers
1. .486
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ilarity 1.9%; Score 26; DB 6; Length 486;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GTGATGTTAGACTTTGAAATTC 26
GTGATGTTAGACTTTGAAATTC 425

8082 487 bp DNA linear PAT 14-JUN-2002
ence 729 from Patent WO0229103.
8082

8082.1 GI:21440787
sapiens (human)
sapiens
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ires, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
expression profiles in liver cancer
ent: WO 0229103-A 729 11-APR-2002;
3 LOGIC INC (US)
Location/Qualifiers
1. .487
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. AA306121"

1 ATGTCATTGTTAGACTTTGAAATTC 26
427 ATGTCATTGTTAGACTTTGAAATTC 452
Db
RESULT 52
BC060315
LOCUS
Homo sapiens cdna clone IMAGE:3879577, partial cds. PRI 2
BC060315 495 bp mRNA linear
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hom
1 (bases 1 to 495)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Sch
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., B
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsi
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Ustin, T.B., Toshitaki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Pete
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., M
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalhus, D
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
23388257
12477932
2 (bases 1 to 495)
Strausberg, R.
Direct Submission
Submitted (20-OCT-2003) National Institutes of Health, N
Gene Collection (GNC), Cancer Genomics Office, National
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 208
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intram
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin,
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., B
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, I
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Lee
Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McC
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., To
Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wig
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information
through the I.M.A.G.E. Consortium/LLNL at: <http://image>

REMARK
COMMENT

IRAK Plate: 134 Row: c Column: 18
 one was selected for full length sequencing because it
 the following selection criteria: Hexamer frequency ORF

Location/Qualifiers

1. .495
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3879577"
 /tissue_type="Lung, large cell carcinoma"
 /clone_lib="NIH MGC 68"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

1.9%; Score 26; DB 9; Length 495;
 100.0%; Pred. No. 0.022;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;

ATGTTAGACTTTGAAATTC 26
 |||||
 ATGTTAGACTTTGAAATTC 424

25 499 bp DNA PAT 18-DEC-2003
 1162 from patent US 6639063.

25.1 GI:40168635

1.
 sified.
 ses 1 to 499)
 3, J.B.D.M., Jobert, S. and Giordano, J.Y.
 and encoded human proteins
 : US 6639063-A 1162 28-OCT-2003;
 Location/Qualifiers

1. .499
 /organism="unknown"
 /mol_type="genomic DNA"

1.9%; Score 26; DB 6; Length 499;
 100.0%; Pred. No. 0.022;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;

CATTGTTAGACTTTGAAATTC 26
 |||||
 CATTGTTAGACTTTGAAATTC 498

78 499 bp DNA PAT 18-SEP-2002
 3 encoded human protein.

78.1 GI:23203896
 2010789-A/1155.
 apiens (human)

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ses 1 to 499)
 s, J.B.D.M., Jobert, S. and Giordano, J.B.
 d encoded human protein
 : JP 2002010789-A 1155 15-JAN-2002;
 CORP

omo sapiens (human)
 P 2002010789-A/1155
 5-JAN-2002

PF 07-AUG-2000 JP 2000280389
 PR 05-AUG-1999 US 60/147499
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN
 GIORDANO

PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N
 C12N1/21,
 PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5
 C12N15/00
 CC EST and encoded human protein
 PH Key Location/Qualifiers
 FT CDS 338..499.

FEATURES
 source
 1. .499
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.9%; Score 26; DB 6; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; G

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 473 ATGTCATTGTTAGACTTTGAAATTC 498

RESULT 55

AX885407 516 bp DNA PAT 18
 LOCUS
 DEFINITION Sequence 1270 from Patent EP1033401.

AX885407

AX885407

VERSION AX885407.1 GI:40041537

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.

TITLE Expressed sequence tags and encoded human proteins

JOURNAL Patent: EP 1033401-A 1270 06-SEP-2000;

Genet (FR)

FEATURES

source
 1. .516
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 343..>516
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAE99545.1"
 /db_xref="GI:40041538"
 /translation="MEYERSPLERKGTLCFHESEDLAEFPVQGYCIHS
 RMSLDLFEISNCLP"

ORIGIN

Query Match 1.9%; Score 26; DB 6; Length 516;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; G

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 478 ATGTCATTGTTAGACTTTGAAATTC 503

RESULT 56

BD025017 516 bp DNA PAT 27
 LOCUS
 DEFINITION Sequence tag and encoded human protein.

BD025017

ACCESSION BD025017

VERSION BD025017.1 GI:22566240

KEYWORDS JP 2001269182-A/1263.

06:25:14 2004

us-09-245-198a-3.oligo.rge

```
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 516)
is,J.B.D.M., Duclair,E. and Jordan,J.Y.
ce tag and encoded human protein
: JP 2001269182-A 1263 02-OCT-2001;
:
omo sapiens (human)
JP 2001269182-A/1263
02-OCT-2001
24-FEB-2000 JP 2000118773
26-FEB-1999 US 60/122487
JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
JORDAN
12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
3/10,
12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
315/40
key Location/Qualifiers
CDS 343..516.
1..516
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1.9%; Score 26; DB 6; Length 516;
larity 100.0%; Pred.No.0.022;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
TCATTGTTAGACTTTGAAATTC 26
|||||
TCATTGTTAGACTTTGAAATTC 503
|||||

934 522 bp DNA linear PAT 18-DEC-2003
nce 3571 from patent US 6639063.
934
934.1 GI:40171044
wn.
wn.
ssified.
ases 1 to 522)
ds,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
and encoded human proteins
t: US 6639063-A 3571 28-OCT-2003;
Location/Qualifiers
1..522
/organism="unknown"
/mol_type="genomic DNA"
1.9%; Score 26; DB 6; Length 522;
larity 100.0%; Pred.No.0.022;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
TCATTGTTAGACTTTGAAATTC 26
|||||
TCATTGTTAGACTTTGAAATTC 501
|||||

.487 522 bp DNA linear PAT 18-SEP-2002
ind and encoded human protein.
.487
```

```
VERSION BD111487.1 GI:23206305
KEYWORDS JP 2002010789-A/3564.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel.
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 522)
Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
EST and encoded human protein
Patent: JP 2002010789-A 3564 15-JAN-2002;
GENSET CORP
OS Homo sapiens (human)
PN JP 2002010789-A/3564
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12.
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT CDS 341..517.
1..522
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1.9%; Score 26; DB 6; Length 522;
Query Match 100.0%; Pred.No.0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0;
QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
|||||
DB 476 ATGTCATTGTTAGACTTTGAAATTC 501
|||||

RESULT 59
AX381939 531 bp DNA linear PAT 1
LOCUS
DEFINITION Sequence 877 from Patent WO0212280.
ACCESSION AX381939
VERSION AX381939.1 GI:19576761
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
1
Pyle,R.A., Xu,J. and Secrist,H.
Compositions and methods for the therapy and diagnosis of
cancer
Patent: WO 0212280-A 877 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1..531
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1.9%; Score 26; DB 6; Length 531;
Query Match 100.0%; Pred.No.0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0;
QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
|||||
DB 459 ATGTCATTGTTAGACTTTGAAATTC 484
|||||
```

22 ap1ens cDNA clone IMAGE:3839189, partial cds.
22 22.1 GI:38328364

apiens (human)

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ber,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

er,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

ul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

S.R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

enko,L., Maruina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

ton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

z,T.E., Brownstein,M.J., Usdin,T.B., Toohiyuki,S.,

ci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

on,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,

an,K.J., Malek,J.A., Gumaratne,P.H., Richards,S.,

,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

on,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,

z,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

rd,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

rd,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

field,Y.S., Krzywinska,M.I., Skalska,U., Smalusz,D.E.,

ch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

tion and initial analysis of more than 15,000 full-length

and mouse cDNA sequences

Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

57

32

ber,G.R.

Submission

ted (13-NOV-2003) National Institutes of Health, Mammalian

ollection (MGC), Cancer Genomics Office, National Cancer

ute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

C Project URL: <http://mgc.nci.nih.gov>

t: MGC help desk

cgapbs@mail.nih.gov

Procurement: ATCC

ibrary Preparation: CLONTECH Laboratories, Inc.

ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)

quencing by: Genome Sequence Centre,

er Agency, Vancouver, BC, Canada

cgsc.bc.ca

Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

a Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

ia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

e, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

l Duane Smalusz, Michael Smith, Lorraine Spence, Jeff Stott,

l Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

Yang, Scott Zuyderduyn, Marco Marra.

distribution: MGC clone distribution information can be found

h the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

IRAL Plate: 51 Row: c Column: 21

lone was selected for full length sequencing because it

the following selection criteria: Hexamer frequency ORF

is.

Location/Qualifiers

1. 534

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3839189"

/tissue_type="Brain, glioblastoma"

/clone_lib="NIH MGC_57"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

ORIGIN

Query Match 1.9%; Score 26; DB 9; Length 534;

Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 26; Conservative 0; Mismatches 0; Indels 0; C

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 438 ATGTCATTGTTAGACTTTGAAATTC 463

RESULT 61

AR415931

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE

1 (bases 1 to 538)

AUTHORS

Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.

TITLE

EST's and encoded human proteins

JOURNAL

Patent: US 6639063-A 3568 28-OCT-2003;

FEATURES

Location/Qualifiers

source

1. 538

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 1.9%; Score 26; DB 6; Length 538;

Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 26; Conservative 0; Mismatches 0; Indels 0; C

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 451 ATGTCATTGTTAGACTTTGAAATTC 476

RESULT 62

BD111484

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 538)

AUTHORS

Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.

TITLE

EST and encoded human protein

JOURNAL

Patent: JP 2002010789-A 3561 15-JAN-2002;

COMMENT

GENSET CORP

OS Homo sapiens (human)

PN JP 2002010789-A/3561

PD 15-JAN-2002

PF 07-AUG-2000 JP 2000280989

PR 05-AUG-1999 US 60/147499

PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN

GIORDANO

PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N

C12N1/21,

PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12NE

C12N15/00

CC EST and encoded human protein

FH Key

Location/Qualifiers

```

CDS      316..492.
          Location/Qualifiers
            1..538
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"

          1.9%; Score 26; DB 6; Length 538;
          larity 100.0%; Pred. No. 0.022;
          Conservative 0; Mismatches 0; Indels 0; Gaps 0;

          TCATTGTTAGACTTTGAAATTC 26
          |||||||
          TCATTGTTAGACTTTGAAATTC 476

          1.9%; Score 26; DB 6; Length 538;
          larity 100.0%; Pred. No. 0.022;
          Conservative 0; Mismatches 0; Indels 0; Gaps 0;

          406 nce 1269 from Patent EP1033401.
          406 nce 1269 from Patent EP1033401.
          406.1 GI:40041535
          sapiens (human)
          sapiens
          yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          lia; Euthera; Primates; Catarrhini; Hominidae; Homo.

          Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.
          ssed sequence tags and encoded human proteins
          t: EP 1033401-A 1269 06-SEP-2000;
          t (FR)

          Location/Qualifiers
            1..540
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
              367..>540
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CA599544.1"
              /db_xref="GI:40041536"
              /translation="MEYERSPLERKQTLCFHSEDLAEPVPGYCIHSLKGCATS
              RMSLLDFEISNCLP"

          1.9%; Score 26; DB 6; Length 540;
          larity 100.0%; Pred. No. 0.022;
          Conservative 0; Mismatches 0; Indels 0; Gaps 0;

          TCATTGTTAGACTTTGAAATTC 26
          |||||||
          TCATTGTTAGACTTTGAAATTC 527

          016 nce tag and encoded human protein.
          016 nce tag and encoded human protein.
          016.1 GI:22566239
          01269182-A/1262.
          sapiens (human)
          sapiens
          yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          lia; Euthera; Primates; Catarrhini; Hominidae; Homo.
          ases 1 to 540)
          ds, J.B.D.M., Duclair, E. and Jordan, J.Y.
          nce tag and encoded human protein
          t: JP 2001269182-A 1262 02-OCT-2001;
          T
          Homo sapiens (human)

```

```

PN JP 2001269182-A/1262
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JE
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C1
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS 367..540.
          Location/Qualifiers
            1..540
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"

          Query Match 1.9%; Score 26; DB 6; Length 540;
          Best Local Similarity 100.0%; Pred. No. 0.022;
          Matches 26; Conservative 0; Mismatches 0; Indels 0;

          QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
          |||||||
          Db 502 ATGTCATTGTTAGACTTTGAAATTC 527

          RESULT 65
          BD224094
          LOCUS 540 bp DNA linear PAT 1
          DEFINITION Sequence characteristic to gene transcription controlled
          hypoxia.
          ACCESSION BD224094
          VERSION BD224094.1 GI:33033864
          KEYWORDS JP 2002525081-A/9.
          SOURCE Homo sapiens (human)
          ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
          Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo
          Binat, P., Skalliter, R. and Feinstein, E.
          TITLE Sequence characteristic to gene transcription controlled
          JOURNAL Patent: JP 2002525081-A 9 13-AUG-2002;
          QUARK BIOTECH INC
          COMMENT OS Homo sapiens (human)
          EN JP 2002525081-A/9
          ED 13-AUG-2002
          PF 27-AUG-1999 JP 2000571058
          PR 27-AUG-1998 US 60/098158, 05-MAY-1999 US 60/132
          PAZ EINAT, RAMI SKALITER, ELENA FEINSTEIN
          PC C12N15/09, A61K31/711, A61K45/00, A61K48/00, A61P9/00, A6
          A61P43/00,
          PC C12Q1/68, G01N33/53, G01N33/566, G01N33/58, G01N33/68, G0
          PC C12N15/00
          CC Sequence characteristic to gene transcription contro
          hypoxia
          FH Key Location/Qualifiers
          FT source 1..540
            Location/Qualifiers
              1..540
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

          Query Match 1.9%; Score 26; DB 6; Length 540;
          Best Local Similarity 100.0%; Pred. No. 0.022;
          Matches 26; Conservative 0; Mismatches 0; Indels 0;

          QY 1 ATGTCATTGTTAGACTTTGAAATTC 26

```

|||||
 CATTGTTAGACTTTGAAATTC 512

35 ce 3572 from patent US 6639063. linear PAT 18-DEC-2003

35 35.1 GI:40171045

n.

n. sified.

ses 1 to 555)

S,J.-B.D.M., Jobert,S. and Giordano,J.-Y.

and encoded human proteins

: US 6639063-A 3572 28-OCT-2003;

Location/Qualifiers

1..555

/organism="unknown"

/mol_type="genomic DNA"

arity 1.9%; Score 26; DB 6; Length 555;

conservative 0; Mismatches 0; Indels 0; Gaps 0;

CATTGTTAGACTTTGAAATTC 26

|||||
 CATTGTTAGACTTTGAAATTC 498

88 d encoded human protein. linear PAT 18-SEP-2002

88

88.1 GI:23206306

2010789-A/3565.

apiens (human)

apiens

ia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ota; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ses 1 to 555)

S,J.-B.D.M., Jobert,S. and Giordano,J.-B.

d encoded human protein

: JP 2002010789-A 3565 15-JAN-2002;

CORP

omo sapiens (human)

P 2002010789-A/3565

5-JAN-2002

7-AUG-2000 JP 2000280989

5-AUG-1999 US 60/147499

EAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI

ANO

12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

/21, 12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC

/00

=a, g, c or t

Location/Qualifiers

338..514

DS

disc feature 541

Location/Qualifiers

1..555

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

arity 1.9%; Score 26; DB 6; Length 555;

conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; C

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

|||||

Db 473 ATGTCATTGTTAGACTTTGAAATTC 498

RESULT 68

BC042807

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) &

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amgcbcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louise

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanan

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information car

through the I.M.A.G.E. Consortium/LLNL at: <http://image.ll>

Series: IRAC Plate: 91 Row: g Column: 13.

Location/Qualifiers

1..560

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5285034"

/tissue_type="Brain, hippocampus"

/clone_lib="NIH_MGC_95"

/lab_host="DH10B"

/note="Vector: pBluescript"

ORIGIN

Query Match 1.9%; Score 26; DB 9; Length 560;

Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 26; Conservative 0; Mismatches 0; Indels 0; C

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

|||||

Db 478 ATGTCATTGTTAGACTTTGAAATTC 503

RESULT 69

AR415932

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AR415932

Sequence 3569 from patent US 6639063.

AR415932

GI:40171042

Unknown.

```

4n.
ssified.
ases 1 to 595)
is,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
and encoded human proteins
: US 6639063-A 3569 28-OCT-2003;
Location/Qualifiers
1. .595
/organism="unknown"
/mol_type="genomic DNA"

1.9%; Score 26; DB 6; Length 595;
larity 100.0%; Pred.No.0.022;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TCATTGTAGACTTTGAAATTC 26
|||||
TCATTGTAGACTTTGAAATTC 538

485
nd encoded human protein.
485
485.1 GI:23206303
02010789-A/3562.
sapiens (human)
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ases 1 to 595)
ds,J.B.D.M., Jobert,S. and Giordano,J.E.
nd encoded human protein
t: JP 2002010789-A 3562 15-JAN-2002;
T CORP
Homo sapiens (human)
JP 2002010789-A/3562
15-JAN-2002
07-AUG-2000 JP 2000280989
05-AUG-1999 US 60/147499
JEAN RAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE PI
DANO
C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,PC
1/21
C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00,PC
5/00
EST and encoded human protein
Key Location/Qualifiers
378. .554.
CDS Location/Qualifiers
1. .595
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

1.9%; Score 26; DB 6; Length 595;
larity 100.0%; Pred.No.0.022;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TCATTGTAGACTTTGAAATTC 26
|||||
TCATTGTAGACTTTGAAATTC 538

060
sapiens cdna FLJ26550 fig, clone LNF01586, highly similar to
aldolase (EC 2.2.1.21).
060

```

```

VERSION AK130060.1 GI:34526798
KEYWORDS oligo capping; fig (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura
Fukuzumi,Y., Fujimori,Y., Komiya,M., Suzuki,Y., Hata,H.
Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyam
Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
2 Unpublished
2 (bases 1 to 1822)
Sugano S. and Suzuki,Y.
Direct Submission
Submitted (31-JUL-2003) Sumio Sugano, Institute of Medica
University of Tokyo, Laboratory of Genome Structure, Huma
Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
(E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry
Economy, Trade and Industry of Japan; cDNA full insert se
Research Association for Biotechnology (RAB); cDNA librar
construction and 5'-end one pass sequencing: Institute of
Science, University of Tokyo, Laboratory of Genome Struct
Genome Center; 3'-end one pass sequencing: RAB; clone sel
full insert sequencing: RAB and Helix Research Institute.
FEATURES
source
1. .1822
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LNF01586"
/cell_type="fibroblast"
/tissue_type="lung"
/clone_lib="INF"
/notes="cloning vector: pME18SFL3"
ORIGIN
Query Match 1.9%; Score 26; DB 9; Length 1822;
Best Local Similarity 100.0%; Pred.No.0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0;
QY 1 ATGTCATTGTAGACTTTGAAATTC 26
Db 1756 ATGTCATTGTAGACTTTGAAATTC 1781
RESULT 72
LOCUS HSM802687 2183 bp mRNA linear PRI 1
DEFINITION Homo sapiens mRNA; cDNA DKFZp762B195 (from clone DKFZp762
ACCESSION AL359585
VERSION AL359585.1 GI:8655645
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homc
1 (bases 1 to 2183)
Bloeker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. a
Wiemann,S.
Direct Submission
Submitted (15-JUN-2000) MIPS, Am Klopferspitz 18a, D-8215
Martinsried, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.c
sequenced by GGF (National Research Centre for Biotechnol
Braunschweig/Germany) within the cDNA sequencing consorti
German Genome Project.
This clone (DKFZp762B195) is available at the RZPD in Ber

```

contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at: <http://www.mips.biochem.mpg.de/proj/cDNA/>.

Location/Qualifiers
1. .2183
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/map="20q13.2-13.2"
/clone="DKFZp762B195"
/tissue_type="melanoma (MeWo cell line)"
/clone_lib="762 (synonym: hmel2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
2140. .2145
2152

1.9%; Score 26; DB 9; Length 2183;
Identity 100.0%; Pred. No. 0.025;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ATTGTTAGACTTTGAAATTC 26
|||||
ATTGTTAGACTTTGAAATTC 2116

75 2650 bp mRNA linear PRI 09-SEP-2003
75 cDNA FLJ42181 f1s, clone THYMU2031368.

75.1 GI:34529902
cloning; f1s (full insert sequence).
sapiens (human)

Meta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Primates; Catarrhini; Hominidae; Homo.

A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
A., Muraashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,
I., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,
Ito, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,
C., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A.,
M.B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K.,
Nagai, T.

Human cDNA sequencing project

1. .2650

T. and Yamamoto, J.

Submission

15-JUL-2003 Takao Isegai, FLJ Project (HRI Team); 2-6-7
-Kamata, Kisarazu, Chiba 292-0818, Japan

l:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

Human cDNA sequencing project supported by Ministry of
Health, Trade and Industry of Japan; cDNA full insert sequencing:

Association for Biotechnology (RAB); cDNA library
technology: Helix Research Institute (HRI) (supported by Japan
Science and Technology Agency (JST) 5'- & 3'-end one pass sequencing: RAB,
National Institute of Advanced Industrial Science and Technology and
National Institute of Health and Biomedical Sciences; HRI and
RAB.

Location/Qualifiers

1. .2650
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="THYMU2031368"
/tissue_type="thymus"
/clone_lib="THYMU2"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 1.9%; Score 26; DB 9; Length 2650;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G

QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
|||||
Db 2584 ATGTCATTGTTAGACTTTGAAATTC 2609

RESULT 74

HSDJ686N3

LOCUS

DEFINITION

HSDJ686N3 110293 bp DNA linear PRI 06
Human DNA sequence from clone RP4-686N3 on chromosome 20q1
contains the 3' part of the gene for a novel ATP dependent
helicase (contains conserved C-terminal helicase domains a
DEAD/DEAH boxes), the KIAA1404 gene, a putative novel gene
STS, GSSs and two CpG islands, complete sequence.

AL049766

AL049766.14 GI:5763746

HTG; CpG island; DEAD box; DEAH box; KIAA1404; RNA helicase

LOCUS Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 110293)

Corby, N.

Direct Submission

Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridges

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Aug 24, 1999 this sequence version replaced gi:5730221.

During sequence assembly data is compared from overlapping

where differences are found these are annotated as variati

variation annotation may not be found in the sequence subm

corresponding to the overlapping clone, as we submit sequ

only a small overlap as described above.

The following abbreviations are used to associate primary

numbers given in the feature table with their source datab

Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Info

on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep This se

the entire insert of clone RP4-686N3 this sequence was fin

follows unless otherwise noted: all regions were either

double-stranded or sequenced with an alternate chemistry o

by high quality data (i.e., phred quality >= 30); an attem

made to resolve all sequencing problems, such as compressi

repeats; all regions were covered by at least one plasmid

or more than one M13 subclone; and the assembly was confir

restriction digest. This sequence was generated from part

bacterial clone contigs of human chromosome 20, constructe

Sanger Centre Chromosome 20 Mapping Group. Further inform

be found at <http://www.sanger.ac.uk/HGP/Chr20>

RP4-686N3 is from the library RPCI-4 constructed by the gr

Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: PCYPAC2.

Location/Qualifiers

1. .110293
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="q13.2-13.2"
/chromosome="20"
/clone="RP4-686N3"
/clone_lib="RPCI-4"

repeat_region 1. .61

/note="AluY repeat: matches 2. .62 of consensus"

207. .21327

/gene="HSDJ686N3.1"

join(<207. .266,518. .683,2195. .2241,2363. .2449

5965. .6138,7449. .7599,10556. .10655,10818. .10959,
12185. .12338,13400. .13481,13583. .13760,16199. .16305,
16486. .16588,19143. .19237,19333. .19456,19844. .19931,
20879. .20953,21066. .21327)
/gene="dJ686N3.1"
/product="dJ686N3.1 (novel ATP dependent RNA helicase
(contains conserved C-terminal helicase domains and
DEAD/DEAH boxes))"
/evidence=not experimental
join(-207. .266,518. .683,2195. .2241,2363. .2449,3680. .3785,
5965. .6138,7449. .7599,10556. .10655,10818. .10959,
12185. .12338,13400. .13481,13583. .13760,16199. .16305,
16486. .16588,19143. .19237,19333. .19456,19844. .19931,
20879. .20953,21066. .21084)
/gene="dJ686N3.1"
/notes="continues from bA120111 (AL357560)
match: Proteins: Tr:O23506 Tr:Q9Z152 Tr:Q9ZCQ0 Wp:CE00337
Tr:Q9ZRZ8 Sw:P25888 Sw:P32892 Sw:Q03903 Tr:O26305
Tr:O87951 Wp:CE22901 Tr:Q07886 Tr:O07897 Tr:O25029
Sw:P38712 Tr:Q9ZM18 Tr:Q9Z6C9 Sw:P21693 Sw:P44701"
/codon_start=1
/evidence=not experimental
/product="dJ586N3.1 (novel ATP dependent RNA helicase
(contains conserved C-terminal helicase domains and
DEAD/DEAH boxes))"
/db_xref="GI:13274192"
/db_xref="GOA:Q96G07"
/db_xref="SWISS-PROT:Q96G07"
/translation="RAATLDEKIEKVRKRTKEDKAGSKLEKEKEGESEPKQ
EDLNDEDESEDEASETSSADENILTKADTKVKDRKKKKKGAGGFGDASQ
YDNLSPQDMNLSRPLIKAITANGFKQPTPIQACIPVGLLGDICACATGTGTAA
FALPVLRIYKRPAPVTVLVTRELGIQVHSVTRQLAQFCNITTCVAGGLDV
KSOEALRAAPDILITPGLRIHLHNCPSFLHSLEVLILDEADRLDEYEPQOMKE
IIRMSCHRTMLFSATDTDEVDKLSVSKNPRVFNSTNTVAPLQCEFRIRPN
REDGRAIVALLTRFTDHWMLPTQTKQAHMHLILGLMLQVGLHGNLSQTORL
EALRFADGIDILTRIVAAAGLDIEGVATVINFMTPTIKHYHVRVGTAKAGAG
RSVLSGCEDEKMLKIVKAAPVKARILPDVILKFRDKIEMKEDVYAVILQLEAG
EKMQQSEAGINTAKLELKEAGVQEPERSWFTQKEEKKEIKALQEFDLALRG
KKRKFKMDKAKGEMTADERSQFILLKAQFAELAKRNRAKRAMPEEPVRG
PAKKQKQGSVPDEBLTNTSKALKQYRAGPSFEERKQLGLPHQRGGNFKSKRYK
RRK"

849. .1146
/notes="AluSx repeat: matches 1. .298 of consensus"
1147. .1423
/notes="AluJo repeat: matches 4. .282 of consensus"
1605. .1912
/notes="AluSx repeat: matches 1. .307 of consensus"
complement(1913. .2334)
/notes="match: GSS: Em:AQ349240"
1952. .2099
/notes="MIR repeat: matches 1. .170 of consensus"
2473. .2589
/notes="FLAM_C repeat: matches 2. .125 of consensus"
2617. .2703
/notes="L2 repeat: matches 2051. .2135 of consensus"
2704. .3004
/notes="AluY repeat: matches 1. .301 of consensus"
3005. .3572
/notes="L2 repeat: matches 2135. .2708 of consensus"
3908. .4051
/notes="FLAM_C repeat: matches 1. .133 of consensus"
4108. .4411
/notes="AluYa5 repeat: matches 1. .304 of consensus"
AluYa5 repeat: matches 1. .304 of consensus"
4441. .4610
/notes="L1MD3 repeat: matches 7557. .7740 of consensus"
4624. .4917
/notes="AluSx repeat: matches 2. .293 of consensus"
5089. .5388
/notes="AluSg repeat: matches 1. .300 of consensus"
5392. .5685
/notes="AluSg repeat: matches 1. .301 of consensus"

repeat_region
5690. .5809
/note="AluSg/x repeat: matches 14. .133 of conse
AluSg/x repeat: matches 14. .133 of consensus"
6281. .6579
/note="AluY repeat: matches 1. .303 of consensus
6896. .7199
/note="AluSg1 repeat: matches 1. .300 of consens
AluSx repeat: matches 1. .303 of consensus"
7814. .7931
/note="MIR repeat: matches 48. .174 of consensus
8181. .8293
/note="L2 repeat: matches 1829. .1957 of consens
8294. .8596
/note="L2 repeat: matches 2347. .2704 of consens
8870. .9009
/note="L1MC4 repeat: matches 7203. .7351 of Cons
9126. .9436
/note="AluJo repeat: matches 1. .308 of consensu
9437. .9492
/note="AluJo repeat: matches 86. .141 of consens
9495. .9597
/note="L1MC4 repeat: matches 7525. .7624 of Cons
9598. .9731
/note="AluSp repeat: matches 1. .127 of consensu
9732. .10040
/note="AluY repeat: matches 1. .308 of consensu
10041. .10191
/note="AluSp repeat: matches 127. .300 of consen
10192. .10371
/note="L1MC4 repeat: matches 7624. .7814 of cons
11182. .11513
/note="L2 repeat: matches 1988. .2346 of consens
11514. .11813
/note="AluSx repeat: matches 1. .300 of consensu
11814. .11963
/note="L2 repeat: matches 2346. .2489 of consens
12695. .12745
/note="L2 repeat: matches 2648. .2700 of consens
12703. .12773
/note="MIR repeat: matches 178. .250 of consensu
12997. .13292
/note="AluY repeat: matches 1. .296 of consensu
complement(13667. .13816)
/note="match: GSS: Em:AQ413681"
complement(13759. .14174)
/note="match: GSS: Em:AQ252228"
13817. .13960
/note="MIR repeat: matches 8. .155 of consensus"
13961. .14303
/gene="dJ686N3.1"
/note="match: GSS: Em:AQ353896"
13961. .14174
/gene="dJ686N3.1"
/note="match: GSS: Em:B63188"
14175. .14282
/note="AluJ/FLAM repeat: matches 2. .87 of conse
AluJ/FLAM repeat: matches 2. .87 of consensus"
14304. .14596
/note="AluJo repeat: matches 3. .300 of consensu
14744. .15145
/note="L2 repeat: matches 2251. .2749 of consens
15261. .15558
/note="AluSc repeat: matches 9. .306 of consensu
15560. .15717
/note="AluJo repeat: matches 145. .293 of consen
15718. .16016
/note="AluSx repeat: matches 1. .297 of consensu

Query Match 1.9%; Score 26; DB 9; Length 110293;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

QY 1 ATGTCATTGTAGACTTTGAAATTC 26

|||||
ATTGTTAGACTTTGAATTC 66465

13 159020 bp DNA linear HTG 27-APR-2000
apiens chromosome 20 clone RP11-564F22 map 20, WORKING DRAFT
E, 21 unordered pieces.

13 2 GI-7651931
XG_PHASE1; HTGS_DRAFT.
apiens (human)

apiens
ata; Chordata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
es 1 to 159020)
B., Linton, L., Nusbaum, C. and Lander, E.
apiens chromosome 20, clone RP11-564F22
shes 1 to 159020)

B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
n, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
wkiy, L., Boukhalil, B., Brown, A., Burkett, G.,
ano, A., Castie, A., Choepel, Y., Colangelo, M., Collins, S.,
re, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,
re, A., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Herre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
J., J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
L., LaRoque, K., Lamarez, R., Landers, T., Lehocsky, J.,
R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
y, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
C., Olivar, T.M., Oliver, J., Peterson, K., Pierre, N.,
C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Santos, R., Schauer, S., Severy, P., Spencer, B.,
Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
ev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
i., Zainoun, J., Zimmer, A. and Zody, M.

Submission
ed (25-MAR-2000) Whitehead Institute/MIT Center for Genome
h, 320 Charles Street, Cambridge, MA 02141, USA
27, 2000 this sequence version replaced gi:7328901.
eats were identified using RepeatMasker:

A.F.A. & Green, P. (1996-1997)
ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
iter: Whitehead Institute/ MIT Center for Genome Research

ter code: WIBR
site: http://www.seq.wi.mit.edu
ract: sequence submissions@genome.wi.mit.edu

----- Project Information

ter project name: L8824
ter clone name: 564_F_22

----- Summary Statistics

encing vector: M13; M77815; 100% of reads
mistry: Dye-terminator Big Dye; 100% of reads
sembly program: Phrap; version 0.960731
sensus quality: 147720 bases at least Q40
sensus quality: 153269 bases at least Q30
sensus quality: 155491 bases at least Q20

ert size: 171000; agarose-fp
lity coverage: 3.9 in Q20 bases; agarose-fp
lity coverage: 4.2 in Q20 bases; sum-of-contigs

: This is a 'working draft' sequence. It currently
sts of 21 contigs. The true order of the pieces
ot known and their order in this sequence record is
rary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1127: contig of 1127 bp in length
* 1128 1227: gap of 100 bp
* 1228 2315: contig of 1088 bp in length
* 2316 2415: gap of 100 bp
* 2416 4526: contig of 2111 bp in length
* 4527 4626: gap of 100 bp
* 4627 8382: contig of 3756 bp in length
* 8383 8482: gap of 100 bp
* 8483 12056: contig of 3574 bp in length
* 12057 12156: gap of 100 bp
* 12157 14923: contig of 2767 bp in length
* 14924 15023: gap of 100 bp
* 15024 19515: contig of 4492 bp in length
* 19516 19615: gap of 100 bp
* 19616 23997: contig of 4382 bp in length
* 23998 24097: gap of 100 bp
* 24098 27747: contig of 3650 bp in length
* 27748 27847: gap of 100 bp
* 27848 33572: contig of 5625 bp in length
* 33573 40493: contig of 6821 bp in length
* 40494 47374: gap of 100 bp
* 47375 47374: contig of 6781 bp in length
* 47375 53409: contig of 6035 bp in length
* 53410 53509: gap of 100 bp
* 53510 62810: contig of 9301 bp in length
* 62811 71685: contig of 8775 bp in length
* 71686 71786: gap of 100 bp
* 71786 80849: contig of 9064 bp in length
* 80850 80949: gap of 100 bp
* 80950 90631: contig of 9682 bp in length
* 90632 90731: gap of 100 bp
* 90732 98024: contig of 7293 bp in length
* 98025 98124: gap of 100 bp
* 98125 110319: contig of 12195 bp in length
* 110320 110419: gap of 100 bp
* 110420 127337: contig of 16917 bp in length
* 127337 127437: gap of 100 bp
* 127437 159020: contig of 31584 bp in length.

FEATURES

source

1. .159020
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20"
/clone="RP11-564F22"
/clone_lib="RP11-11 Human Male BAC"
misc_feature 1. .1127
/note="assembly_fragment"
misc_feature 1228. .2315
/note="assembly_fragment"
misc_feature 2416. .4526
/note="assembly_fragment"
misc_feature 4627. .8382
/note="assembly_fragment"
misc_feature 8483. .12056
/note="assembly_fragment"
misc_feature 12157. .14923
/note="assembly_fragment"
misc_feature clone_end:17
vector_side:right
misc_feature 15024. .19515
/note="assembly_fragment"
misc_feature 19616. .23997
/note="assembly_fragment"
misc_feature 24098. .27747

06:25:14 2004

us-09-245-198a-3.oligo.rge

```
/note="assembly_fragment"  
27848..33472  
/note="assembly_fragment"  
33573..40393  
/note="assembly_fragment"  
40494..47274  
/note="assembly_fragment"  
47375..53409  
/note="assembly_fragment"  
53510..62810  
/note="assembly_fragment"  
62911..71685  
/note="assembly_fragment"  
71786..80849  
/note="assembly_fragment"  
80950..90631  
/note="assembly_fragment"  
90732..98024  
/note="assembly_fragment"  
clone_end:SP6  
vector_side:right"  
98125..110319  
/note="assembly_fragment"  
110420..127336  
/note="assembly_fragment"  
127437..159020  
/note="assembly_fragment"
```

1.9%; Score 26; DB 2; Length 159020;
arity 100.0%; Pred.No. 0.036;
conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
CCATTGTAGACTTTGAATTTC 26  
|||||  
CCATTGTAGACTTTGAATTTC 93106
```

April 8, 2004, 22:35:57
acs

16:25:16 2004

us-09-245-198a-3.oligo.rst

GenCore version 5.1.6
copyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 8, 2004, 20:38:11 ; Search time 3960 seconds
(without alignments)
10353.731 Million cell updates/sec

-09-245-198a-3

3

tgctcattgtgacttga.....gacaaaattgtgataaatgg 1373

GO_NUC

op 60.0 , Gapext 60.0

13289 seqs, 14931090276 residues

s satisfying chosen parameters: 55026578

th: 0

th: 2000000000

sting first 100 summaries

T:*

em_estba:*

em_esthum:*

em_estin:*

em_estnu:*

em_estov:*

em_estpl:*

em_estro:*

em_htc:*

gb_est1:*

gb_est2:*

gb_htc:*

gb_est3:*

gb_est4:*

gb_est5:*

em_estfun:*

em_estom:*

em_gss_hum:*

em_gss_inv:*

em_gss_pln:*

em_gss_vrt:*

em_gss_fun:*

em_gss_mam:*

em_gss_mus:*

em_gss_pro:*

em_gss_rod:*

em_gss_pbg:*

em_gss_vrl:*

gb_gss1:*

gb_gss2:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry	ch	Length	DB	ID	Description
1.1	776	13	BM9090012	BM9090012	BM9090012
1.8	777	12	BM19200	BM19200	BM19200
1.6	1071	12	BM921213	BM921213	BM921213
1.1	963	13	BM671259	BM671259	BM671259

8.7 639 10 BB642326
 8.7 692 13 BY748962
 8.7 1187 12 B0053284
 8.7 2237 11 AK044387
 7.4 1926 12 BM906056
 7.3 289 10 BE503779
 7.1 254 12 BM128317
 6.9 837 14 CB961075
 6.8 105 10 AW191845
 6.3 538 10 BF821434
 6.1 140 9 A1582027
 5.9 576 9 A1535912
 5.3 278 10 AW417023
 5.3 404 10 BF044430
 5.3 542 10 BF041509
 5.2 257 14 R55285
 5.2 784 12 BG686384
 4.2 571 10 BF073881
 3.8 584 10 AW917574
 3.6 426 10 AW236926
 3.6 561 10 AW763237
 3.4 237 29 CG616340
 3.4 269 29 CG564287

ALIGNMENTS

12 Soares breast 2NBHst mRNA linear EST 23-JAN-2003
 98E15243 ; IMAGE:154742, mRNA sequence.
 12.1 GI:27821952
 apiens (human)
 apiens
 ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ses 1 to 776)
 L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
 f., U., Schneider, D. and Korn, B.
 UnigeneSet - RZPD3
 ished (2003)
 t: Ina Rols
 eutsches Ressourcenzentrum fuer Genomforschung GmbH
 enheimer Feld 580, D-69120 Heidelberg, Germany
 IMAGp998E15243.
 B; I.M.A.G.E. cDNA Clone Collection;
 UnigeneSet - RZPD3 (RZPDLIB NO.972)
 /www.rzpd.de/CloneCards/cgi-
 owLib.pl.cgi/response?libNo=972 Contact: Ina Rols
 eutsches Ressourcenzentrum fuer Genomforschung GmbH
 rweg 6, D-14059 Berlin, Germany
 49 30 32639 101
 49 30 32639 111
 pd.de
 lone is available royalty-free from RZPD;
 RZPD (clone.rzpd.de) for further information. Seq primer:
 Primer sequence: TTTCACACAGAAACACAGCTATGAC.
 Location/Qualifiers
 1..776
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGp998E15243 ; IMAGE:154742"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares breast 2NBHst"
 /note="Organ: breast; Vector: p173D (Pharmacia) with a
 modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) I
 TGTTACCAATCTGAAGTGGAGCGCGCCCTTTTCTTTTCTTTT
 double-stranded cDNA was ligated to Eco RI adapt
 (Pharmacia), digested with Not I and cloned into
 and Eco RI sites of a modified p173 vector (Phar
 Library went through one round of normalization t
 230. Library constructed by Bento Soares and M.F.
 Bonaldo."

ORIGIN

Query Match 52.1%; Score 716; DB 13; Length 776;
 Best Local Similarity 100.0%; Pred.No. 0;
 Matches 716; Conservative 0; Mismatches 0; Indels 0;
 QY 600 TCAGGTGCACATTCATGAGGGGAGGCTGTCTACTGAAGCTGGACTTGTCTGTG
 Db 1 TCAGGTGCACATTCATGAGGGGAGGCTGTCTACTGAAGCTGGACTTGTCTGTG
 QY 660 TGTCTGCGCCCTGCGCTGCGCTGGAGGAATTCAGCCACTCGCGCCAGTTCCTCTC
 Db 61 TGTCTGCGCCCTGCGCTGCGCTGGAGGAATTCAGCCACTCGCGCCAGTTCCTCTC
 QY 720 CCAGCTCGCCCTGCGCTGCGCTGGAGGAATTCAGCCACTCGCGCCAGTTCCTCTC
 Db 121 CCAGCTCGCCCTGCGCTGCGCTGGAGGAATTCAGCCACTCGCGCCAGTTCCTCTC
 QY 780 GATCGCACCTCCCTGCGCCCTGCGCTGGAGGAATTCAGCCACTCGCGCCAGTTCCTCTC
 Db 181 GATCGCACCTCCCTGCGCCCTGCGCTGGAGGAATTCAGCCACTCGCGCCAGTTCCTCTC
 QY 840 CTTCCAGGTTCACTGAGGGGCGCTGGTTCCTCCACAGTCTGCTCCAGGCTCGCGGCT
 Db 241 CTTCCAGGTTCACTGAGGGGCGCTGGTTCCTCCACAGTCTGCTCCAGGCTCGCGGCT
 QY 900 TCGACAGTCTCTGCGGCGCCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT
 Db 301 TCGACAGTCTCTGCGGCGCCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT
 QY 960 ACCTGCGCCCTCCCTCTAGAGGCTGCTGGGCTGCTGACGTTGTTTTCATCCCCACF
 Db 361 ACCTGCGCCCTCCCTCTAGAGGCTGCTGGGCTGCTGACGTTGTTTTCATCCCCACF
 QY 1020 TACAGTATTCCTCACTCTTATCTTACAACTCCCGCCAGCCCACTCTCCACTCACT
 Db 421 TACAGTATTCCTCACTCTTATCTTACAACTCCCGCCAGCCCACTCTCCACTCACT
 QY 1080 CCCCACATCCCTGACCCCTTTCAGGCGCCCGCTGCGGCTGCGGCTGCGGCTGCGGCT
 Db 481 CCCCACATCCCTGACCCCTTTCAGGCGCCCGCTGCGGCTGCGGCTGCGGCTGCGGCT
 QY 1140 CCAGGCAATTGTGTTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAGAGACCCAC
 Db 541 CCAGGCAATTGTGTTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAGAGACCCAC
 QY 1200 GGCACTAAGAGGGGCTGGACCTGCGCGGAGAGCCAAAGAGACTGGGCTTAGGCC
 Db 601 GGCACTAAGAGGGGCTGGACCTGCGCGGAGAGCCAAAGAGACTGGGCTTAGGCC
 QY 1260 GTTCCCAAAATGTGAGGGGCGAGAAACAGCAAGTCTCTCCCTTGAGAAATTCCTG
 Db 661 GTTCCCAAAATGTGAGGGGCGAGAAACAGCAAGTCTCTCCCTTGAGAAATTCCTG

RESULT 2
 BI819200
 LOCUS
 DEFINITION 603034614F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175
 mRNA sequence.
 ACCESSION BI819200
 VERSION BI819200.1 GI:15930750
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 777
http://mhc.nci.nih.gov/
l. Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
cgapb-remail.nih.gov
Procurement: Life Technologies, Inc.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
quencing by: Incyte Genomics, Inc.
Distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LLAM1437 row: 1 column: 03
ality sequence stop: 759.
Location/Qualifiers
1. .777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5175698"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27, and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
45.8%; Score 629; DB 12; Length 777;
rity 99.7%; Pred. No. 3.1e-309;
nservative 0; Mismatches 2; Indels 0; Gaps 0;
AGGAGGACCCAGGACCCCTCGAAGTGAATCCAGAGAGAAAGCCAGATCC 347
|||||
AGGAGGACCCAGGACCCCTCGAAGTGAATCCAGAGAGAAAGCCAGATCC 72
|||||
CTTTCTGAAACCGACTAGTTTCGGCTCGCAGAGTGCACCTTAAGCGCGGAAC 407
|||||
CTTTCTGAAACCGACTAGTTTCGGCTCGCAGAGTGCACCTTAAGCGCGGAAC 132
|||||
CTCGAAGGCGATCGCAGCCCATTTATGAAGTTTCATCCAGCAGCTGGACGACGG 467
|||||
CTCGAAGGCGATCGCAGCCCATTTATGAAGTTTCATCCAGCAGCTGGACGACGG 192
|||||
AGGAGAGTGTGGACGGACAGTGAAGTGGCTGGAGGAGGAGCCAGATCAACAGCTC 527
|||||
AGGAGAGTGTGGACGGACAGTGAAGTGGCTGGAGGAGGAGCCAGATCAACAGCTC 252
|||||
CTCTGCCCTACAAACCGCAGATCGGAGTGTATAGTACCCCGGCTGGGCTCTA 587
|||||
CTCTGCCCTACAAACCGCAGATCGGAGTGTATAGTACCCCGGCTGGGCTCTA 312
|||||
TGACTGTGAGTGACATTTTATGAGGGGAGGCTGTCTACCTGAAGCTGGACTT 647
|||||
TGACTGTGAGTGACATTTTATGAGGGGAGGCTGTCTACCTGAAGCTGGACTT 372
|||||
TGGATGTGTGCTGGCCCTCGCTGCTGGAGGAATTCAGCAGCTGGCCGACG 707
|||||
TGGATGTGTGCTGGCCCTCGCTGCTGGAGGAATTCAGCAGCTGGCCGACG 432
|||||
TCGGGCCCCAGCTCCGCTCTCGCAGGTGCTGGGCTGTGGCCCTCGGCCAGG 767
|||||
TCGGGCCCCAGCTCCGCTCTCGCAGGTGCTGGGCTGTGGCCCTCGGCCAGG 492
|||||

QY 768 GTCTCTCCCTCGGATCCGACCCCTCCCTCGGCCCCATCTCAAGGCTGCCCTTCTCT
|||||
Db 493 GTCTCTCCCTCGGATCCGACCCCTCCCTCGGCCCCATCTCAAGGCTGCCCTTCTCT
|||||
QY 828 CTACTTCGAGCTCTTCCAGGTTCACTGAGGGGCCCTGTCTCCCCACAGTGTCTCCCA
|||||
Db 553 CTACTTCGAGCTCTTCCAGGTTCACTGAGGGGCCCTGTCTCCCCACAGTGTCTCCCA
|||||
QY 888 TGCAGGCTCCCTCGACAGCTCTCTGGGACACCGGTCCTCTGCCCTCTGCCCTCAGC
|||||
Db 613 TGCAGGCTCCCTCGACAGCTCTCTGGGACACCGGTCCTCTGCCCTCAGC
|||||
QY 948 TCTTTGTCTCCAGACTGCCCCCTCCCTCTAGAGGCTGCTGGGCTCTTTCAGCTGTT
|||||
Db 673 TCTTTGTCTCCAGACTGCCCCCTCCCTCTAGAGGCTGCTGGGCTCTTTCAGCTGTT
|||||
QY 1008 ATCCACATAA 1018
|||||
Db 733 ATCCACATAA 743
|||||
RESULT 3
BM921213
LOCUS
DEFINITION BM921213 1071 bp mRNA linear EST 12-
AGENCOURT 6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAGE
5', mRNA sequence.
ACCESSION BM921213
VERSION BM921213.1 GI:19371592
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1071)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (M
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12786 row: p column: 02
High quality sequence stop: 656.
Location/Qualifiers
1. .1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5752561"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroye
source anonymous pool of 6 male brains, age range
male lung, age 27; and 1 male testis, age 69. Lib
oligo-dr primed and directionally cloned (EcoRV si
destroyed upon cloning). Average insert size 1.8
insert size range 1-3 kb. Library is normalized an
enriched for full-length clones and was construct
Gruber (Invitrogen). Research Genetics tracking
021. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 44.6%; Score 613; DB 12; Length 1071;
Best Local Similarity 99.7%; Pred. No. 4.8e-301;
Matches 713; Conservative 0; Mismatches 2; Indels 0; Ga
QY 412 GCTCGAAGAGCGATCGCAGCCCATTTATGAAGTTTCATCCAGCAGCTGGACGACGGA

|||||
GAGAGCGATCGCAGCCCAATATGAAGTTATCCACGACCTGGAGCAGCGAGCG 71
|||||
CAGGTGGAGCGGACAGTAGTGGCTGGGAGGAGCAGATCAACAGCTCCAGC 531
|||||
CAGGTGGAGCGGACAGTAGTGGCTGGGAGGAGCAGATCAACAGCTCCAGC 131
|||||
TGCGCTACAAACCGCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTTACTAC 591
|||||
TGCGCTACAAACCGCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTTACTAC 191
|||||
ACTGTAGGTGCATTTGATAGGGAGAGCTGTCTACCTGAAGCTGGACTTGTG 651
|||||
ACTGTAGGTGCATTTGATAGGGAGAGCTGTCTACCTGAAGCTGGACTTGTG 251
|||||
ATGTGTGTGGGCTGGCTGGCTGGCTGGAGGAATTTCTCAGCACTGGCGGCAGTTCC 711
|||||
ATGTGTGTGGGCTGGCTGGCTGGCTGGAGGAATTTCTCAGCACTGGCGGCAGTTCC 311
|||||
GGCCCCAGCTCCGCTCTGCGCAGGTGTCTGGGCTGTGGGCTGGCGGCCAGGCTCC 771
|||||
GGCCCCAGCTCCGCTCTGCGCAGGTGTCTGGGCTGTGGGCTGGCGGCCAGGCTCC 371
|||||
TGCGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGGCCCCCTTCTCACCCTAC 831
|||||
TGCGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGGCCCCCTTCTCACCCTAC 431
|||||
GACTTCTCCAGGTTCACTGAGGGGCTGGTCTCCCACTGCTCCCGAGGCTGCC 891
|||||
GACTTCTCCAGGTTCACTGAGGGGCTGGTCTCCCACTGCTCCCGAGGCTGCC 491
|||||
CCCTCCGACAGCTCTCTGGGACCCGCTGGCTCTGCCCCACCTCAGCGGCTCTT 951
|||||
CCCTCCGACAGCTCTCTGGGACCCGCTGGCTCTGCCCCACCTCAGCGGCTCTT 551
|||||
CCAGACTGGCCCTCCCTCTAGAGGCTGGCTGGGCTGTTCAGTGTTTCCATCC 1011
|||||
CCAGACTGGCCCTCCCTCTAGAGGCTGGCTGGGCTGTTCAGTGTTTCCATCC 611
|||||
TAAATACAGTATTCCTACTTATCTTACAACTCCCGCCAGCGGCTCCACCT 1071
|||||
TAAATACAGTATTCCTACTTATCTTACAACTCCCGCCAGCGGCTCCACCT 671
|||||
AGCTCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGTCTCGATCTCCCGCC 1126
|||||
AGCTCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGTCTCGATCTCCCGCC 726

59 963 bp mRNA linear EST 15-JUL-2002
JRT 8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
NA sequence.

59
59.1 GI:21782093

apiens (human)

apiens

sta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ses 1 to 963)

C http://mgs.nci.nih.gov/.

al Institutes of Health, Mammalian Gene Collection (MGC)

ished (1999)

t: Robert Strausberg, Ph.D.

cgapb-remail.nih.gov

Procurement: ATCC

Library Preparation: Rubin Laboratory

Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

sequencing by: Agencourt Bioscience Corporation

distribution: MGC clone distribution information can be

through the I.M.A.G.E. Consortium/LNL at:

/image.llnl.gov

Plate: LLCM2456 row: 1 column: 13
High quality sequence stop: 565.
Location/Qualifiers
1. .963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6274716"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/notes="Organ: salivary gland; Vector: pORB7; Site
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using
following 5' adaptor: GGCAGGAG(G). Library const
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-c
synthesis kit (Stratagene) and Superscript II RT
Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 44.1%; Score 605; DB 13; Length 963;
Best Local Similarity 100.0%; Pred. No. 5.7e-297;
Matches 605; Conservative 0; Mismatches 0; Indels 0; G

Qy 408 ACGGGCTCGAAGAGCGATCGCAGCCCAATATGAAGTTATCCACGACCTGGACAGG
Db 39 ACGGGCTCGAAGAGCGATCGCAGCCCAATATGAAGTTATCCACGACCTGGACAGG
Qy 468 AGCGCAGCGAGGTGTGGACGGGACAGTGTGGCTGGGAGGAGCCAGAAATCAACA
Db 99 AGCGCAGCGAGGTGTGGACGGGACAGTGTGGCTGGGAGGAGCCAGAAATCAACA
Qy 528 CAGCCCTCTCGGCTACAAACCGCAGATCGGGAGTGTATAGTACACCGGGCTGGGC
Db 159 CAGCCCTCTCGGCTACAAACCGCAGATCGGGAGTGTATAGTACACCGGGCTGGGC
Qy 588 CTACCTGTACTGTCTCAGGTGCACCTTTGATGAGGGGAGGCTGTCTACCTGAAGCTGG
Db 219 CTACCTGTACTGTCTCAGGTGCACCTTTGATGAGGGGAGGCTGTCTACCTGAAGCTGG
Qy 648 GCTGGTGGATGGTGTGTCTGGCCCTGCGCTGCGCTGGAGGAATTTCTCAGCCACTGGCG
Db 279 GCTGGTGGATGGTGTGTCTGGCCCTGCGCTGCGCTGGAGGAATTTCTCAGCCACTGGCG
Qy 708 TTCCCTCGGGCCCCAGCTCCGCTCTGCGAGGTGTCTGGGCTGTGGGCTGGCGGC
Db 339 TTCCCTCGGGCCCCAGCTCCGCTCTGCGAGGTGTCTGGGCTGTGGGCTGGCGGC
Qy 768 GTCTCTCCCTCGGATCGCAGCCCTCCCTCGGGGCCATCTCAAGGCTGCCCTTTCC
Db 399 GTCTCTCCCTCGGATCGCAGCCCTCCCTCGGGGCCATCTCAAGGCTGCCCTTTCC
Qy 828 CTACTTCGAGCTCTTCTCAGGTTCACGTAGGGGGCCCTGGTCTCCGCCACAGTCGTCGCC
Db 459 CTACTTCGAGCTCTTCTCAGGTTCACGTAGGGGGCCCTGGTCTCCGCCACAGTCGTCGCC
Qy 888 TGGCGGCTCCCTCGCAGCTCTCTGGGACCCCGGTCCTGCCCCACCCCTCAG
Db 519 TGGCGGCTCCCTCGCAGCTCTCTGGGACCCCGGTCCTGCCCCACCCCTCAG
Qy 948 TCTTTGTCTCCAGACCTGCCCTCCCTCTAGAGGCTGCTGGGCTGTTCACAGTGT
Db 579 TCTTTGTCTCCAGACCTGCCCTCCCTCTAGAGGCTGCTGGGCTGTTCACAGTGT
Qy 1008 ATCCC 1012
Db 639 ATCCC 643

RESULT 5

CA396679

LOCUS

CA396679

567 bp

mRNA

linear

EST 06

.y1 Human Retinal pigment epithelium/choroid cDNA
 malized, unamplified): cs Homo sapiens cDNA clone cs80h07
 A sequence.

9.1 GI:24733297

piens (human)

ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Primates; Catarrhini; Hominidae; Homo.

es 1 to 567)

G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,

n,J.W., Bouffard,G., Smith,D. and Peterson,K.

ed sequence tag analysis of human RPE/choroid for the

Project: Over 6000 non-redundant transcripts, novel genes

ice variants

s. 8 (4), 205-220 (2002)

0

0

on Molecular Structure and Function

1 Eye Institute

NIH, Bethesda, MD 20892-2740, USA

1 402 3452

1 496 0078

graeme@helix.nih.gov

80 row: h column: 07

mer: M13RPI reverse primer (ABI).

Location/Qualifiers

1. .567

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="cs80h07"

/tissue_type="RPE/choroid"

/dev_stage="Adult"

/lab_host="EMDH10B"

/clone_lib="Human Retinal pigment epithelium/choroid cDNA

(Un-normalized, unamplified): cs"

/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor

eyes (75-80 years old) yielded approximately 600 mg of

dissected RPE/choroid tissue. This in turn yielded 340 ug

of total RNA and 7 ug of mRNA. A directionally cloned cDNA

library in the pCMVSPORT6 vector was constructed at Life

Technologies (Rockville, MD; now part of Invitrogen Corp),

essentially following the protocols of the SuperScript

Plasmid System (Invitrogen Corp).

<http://www.invitrogen.com/>). The library code

designation was cs. For this library, cDNA inserts were

cloned into the NotI/MluI sites of the vector. EST

analysis was performed on the unamplified library at the

NIH Intramural Sequencing Center (NISC)."

41.3%; Score 567; DB 14; Length 567;

ity 100.0%; Pred. No. 1.3e-277;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

3TCCCGGATGGGGGGGGGTGAGGACAGGACAGCCCCCGCCCGATGCGCGCC 114

3TCCCGGATGGGGGGGGGGTGGAGGACAGGACAGCCCCCGCCCGATGCGCGCC 60

GAGCCAGAGCGGAGGGGGGGCGCGGGGGGAGCGGGCACCGCCCTGCTGTCGGC 174

GAGCCAGAGCGGAGGGGGGGCGCGGGGGGAGCGGGCACCGCCCTGCTGTCGGC 120

GCTGGGCGCTGGGCGCTGGCGCTGCGCTCTCGGCTCTGCTGCGCGCGTGTGTCAGT 234

GCTGGGCGCTGGGCGCTGGCGCTGCGCTCTCGGCTCTGCTGCGCGCGTGTGTCAGT 180

GAGCCGGGCGATCGTGTGCGCGCGGAGGCGCTGCGCGAGGAGCTGTGTCAGAG 294

GAGCCGGGCGATCGTGTGCGCGCGGAGGCGCTGCGCGAGGAGCTGTGTCAGAG 240

QY 295 GAGGACACGAGCCCTCGAACTGAATCCCGACAGAGAAAGCCAGGATCCTGCG
 DB 241 GAGGACACGAGCCCTCGAACTGAATCCCGACAGAGAAAGCCAGGATCCTGCG
 QY 355 TTCCTGAACCGACTAGTTCGGCCTCGCAGAACTGCACTTAAAGCGCGGAAAAACACG
 DB 301 TTCCTGAACCGACTAGTTCGGCCTCGCAGAACTGCACTTAAAGCGCGGAAAAACACG
 QY 415 CGAAGAGCGATCGGAGCCCATATGAGTTTCATCCAGCAGCTTGGACAGCGAGCGGCG
 DB 361 CGAAGAGCGATCGGAGCCCATATGAGTTTCATCCAGCAGCTTGGACAGCGAGCGGCG
 QY 475 GCAGGTGTGGACGGGACAGTGTAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC
 DB 421 GCAGGTGTGGACGGGACAGTGTAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC
 QY 535 CTGCGCTACAAACCGCCAGATCGGGGAGTTTATGATCAACCGCGCTGGGCTTACTAC
 DB 481 CTGCGCTACAAACCGCCAGATCGGGGAGTTTATGATCAACCGCGCTGGGCTTACTAC
 QY 595 TACTGTCAAGTGCACCTTTGATGAGGGG 621
 DB 541 TACTGTCAAGTGCACCTTTGATGAGGGG 567

RESULT 6

LOCUS

CB141389

DEFINITION

K-EST0194999 L15CKK1 Homo sapiens cDNA clone L15CKK1-30-E06

mRNA sequence.

ACCESSION

CB141389

VERSION

CB141389.1 GI:28116436

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., K

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

CONTACT: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kr.ibm.re.kr

Plate: 30 row: E column: 06

High quality sequence stop: 545.

Location/Qualifiers

1. .545

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="L15CKK1-30-E06"

/sex="M"

/cell_line="CK-K1"

/lab_host="Top10P"

/clone_lib="L15CKK1"

/note="Organ: Liver; Vector: pcNS-D2; Site: 1; EcoR

Site 2: NotI; The poly (A) + RNA was dephosphorylat

bacterial alkaline phosphatase (BAP) and then deca

with tabacco acid pyrophosphatase (TAP). The decap

intact mRNA was ligated with DNA-RNA linker includ

EcoRI site by treatment of T4 RNA ligase and the f

strand cDNA was synthesized from oligo dt-selected

priming with dt-tailed vector. The dt-tailed vecto

adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digesti

EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells *E. coli* Top10[®] by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

39.7%; Score 545; DB 14; Length 545;
arity 100.0%; Pred. No. 2.1e-266;
conservative 0; Mismatches 0; Indels 0; Gaps 0;

GAACTGAATCCCGACAGACAGAGAAAGCCAGGATCCTCGGCTTTCCTGAACCGACT 368
GAACTGAATCCCGACAGACAGAGAAAGCCAGGATCCTCGGCTTTCCTGAACCGACT 60
CGGCTCTCGCAGAAAGTGCACTTAAAGGCCGGAACACACCGGCTCGAAGAGCGATCGC 428
CGGCTCTCGCAGAAAGTGCACTTAAAGGCCGGAACACACCGGCTCGAAGAGCGATCGC 120
CAATTATGAAGTTTCATCCAGGACCTGGACAGCAGGAGCGAGGAGCGTGTGGACGG 498
CAATTATGAAGTTTCATCCAGACCTGGACAGCAGGAGCGAGGAGCGTGTGGACGG 180
GTGAGTGGCTGGGAGGAAGCCAGAAATCAACAGCTCCAGCCCTCTGGCTACACCG 548
GTGAGTGGCTGGGAGGAAGCCAGAAATCAACAGCTCCAGCCCTCTGGCTACACCG 240
ATCGGGAGTTTATAGTCAACCGGCTGGGCTCTACTACTGTACTGTCTGAGTGCA 608
ATCGGGAGTTTATAGTCAACCGGCTGGGCTCTACTACTGTACTGTCTGAGTGCA 300
GATGAGGGGAAGGCTGTCTACTGAAGCTTGACCTTGCTGGTGGATGCTGTGCTGGC 668
GATGAGGGGAAGGCTGTCTACTGAAGCTTGACCTTGCTGGTGGATGCTGTGCTGGC 360
CGCTGTGCTTGGAGGAATTCAGACCACTCGCGCCAGTTCCCTCTCGGGCCCCAGTCCG 728
CGCTGTGCTTGGAGGAATTCAGACCACTCGCGCCAGTTCCCTCTCGGGCCCCAGTCCG 420
TGGCAGGTGTCTGGGCTGTGTGCTGGGCCAGAGGTCCTCCCTCGGGATTCGGCAC 788
TGGCAGGTGTCTGGGCTGTGTGCTGGGCCAGAGGTCCTCCCTCGGGATTCGGCAC 480
CCCTGGGGCCCATCTCAAGGCTGCCCCCTTCTCACTACTTCCGACTCTTCCAGGT 848
CCCTGGGGCCCATCTCAAGGCTGCCCCCTTCTCACTACTTCCGACTCTTCCAGGT 540

T 853
T 545

81 828 bp mRNA . linear EST 07-SEP-2001
254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',
sequence

81.1 GI:15489620

apiens (human)
apiens
apiens
ota;
ata; Craniata; Chordata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
ses 1 to 828)
C http://mgc.nci.nih.gov/.
al Institutes of Health, Mammalian Gene Collection (MGC)
ished (1999)
t: Robert Strausberg, Ph.D.
cgabs-remail.nih.gov
Procurement: Miklos Palkovits, M.D., Ph.D.
Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM11722 row: k column: 13
High quality sequence stop: 776.

[illegible]

```

1. 828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5285892"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH MGC 96"
/name="Organ; brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-Xhc
(gcgagg); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTNN-3', size-selected for aver
insert size 2.3 kb and normalized to R01 5. This
primary library enriched for full-length clones
constructed using the Cap-trapper method (Carnine
preparation). Library constructed by M. Brownste
(NIH/NHGRI, National Institutes of Health). Note
NIH MGC Library."

```

ORIGIN

Query Match	37.7%	Score 518	DB 12	Length 828
Best Local Similarity	99.7%	Pred. No. 1.4e-252		
Matches 618	Conservative 0	Mismatches 2	Indels 0	
QY	312	GGAACCTGAATCCCCAGACAGAGAAAGAACAGAGATCCTGGCGCTTTCTGTGAACCGAC		
Db	43	GGAACCTGAATCCCCAGACAGAGAAAGAACAGAGATCCTGGCGCTTTCTGTGAACCGAC		
QY	372	TCGGCCTCGCAGAGTGCACCTTAAGGCCCGGAAAAACACGGGCTCGAAGAGCGATCC		
Db	103	TCGGCCTCGCAGAGTGCACCTTAAGGCCCGGAAAAACACGGGCTCGAAGAGCGATCC		
QY	432	CCATTTATGAAGTTTCATCCAGGACCTTGGACAGAGCGAGCGGAGGAGGTGTGGAC		
Db	163	CCATTTATGAAGTTTCATCCAGGACCTTGGACAGAGCGAGCGGAGGAGGTGTGGAC		
QY	492	AGTCAGTGGCTTGGAGGAAGCCAGAGATCAACAGCTCCAGGCCCTCTGCCTACAAC		
Db	223	ATTGAGTGGCTTGGAGGAAGCCAGAGATCAACAGCTCCAGGCCCTCTGCCTACAAC		
QY	552	GATCGGGAGTTTATAGTACCCGGGCTGGGCTCTACTACCTGTACTGTGTCAAGTGC		
Db	283	GATCGGGAGTTTATAGTACCCGGGCTGGGCTCTACTACCTGTACTGTGTCAAGTGC		
QY	612	TGATGAGGGGAAGGCTGTCTACTCTGAAGCTGTGCACCTTGTGTGTGATGTGTGCTGC		
Db	343	TGATGAGGGGAAGGCTGTCTACTCTGAAGCTGTGCACCTTGTGTGTGATGTGTGCTGC		
QY	672	GCCTGCTCGGAGAAATTCAGCCACTCGGCCAGTTCCCTCGGGCCCCCAGCTCC		
Db	403	GCCTGCTCGGAGAAATTCAGCCACTCGGCCAGTTCCCTCGGGCCCCCAGCTCC		
QY	732	CTGCCAGGTGTCTGGGCTGTGTGGCCCTTGGGGCCAGGTCCTCCCTGCGGATTCGGC		
Db	463	CTGCCAGGTGTCTGGGCTGTGTGGCCCTTGGGGCCAGGTCCTCCCTGCGGATTCGGC		
QY	792	CCCTTGGGGCCCATCTCAAGGCTGCCCCCTTCTCACTACTTTCGAGCTCTTCCAGG		
Db	523	CCCTTGGGGCCCATCTCAAGGCTGCCCCCTTCTCACTACTTTCGAGCTCTTCCAGG		
QY	852	CTGAGGGGGCCCTGTCTCCCCACAGTCGCTCCAGGCTGCCGGCTCCCTCTGACAGC		
Db	583	CTGAGGGGGCCCTGTCTCCCCAGCTGTCTCCAGGCTGCCGGCTCCCTCTGACAGC		

```
ACCGGTCCTCCCTCTG 931
|||||
ACCGGTCCTCCCTCTG 662

3 728 bp mRNA linear EST 11-OCT-2001
41F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405459 5',
|||||
3 quence.

3.1 GI:16044066
piens (human)
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 728)
http://imgc.ncl.nih.gov/.
l Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
cgabbs@mail.nih.gov
Procurement: ATCC
library Preparation: Life Technologies, Inc.
ibrary Arrayed by: The I.M.A.G.E. Consortium (LLNL)
quencing by: Incyte Genomics, Inc.
distribution: MGC clone distribution information can be
hrough the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LLAM12034 row: a column: 12
ality sequence stop: 728.
Location/Qualifiers
1. .728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405459"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

36.9%; Score 506; DB 12; Length 728;
rity 100.0%; Pred. No. 1.8e-246;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

GCCTGCCAGAGGAGCTGTGGCAGAGGAGGACGAGCCGTCGGACTGAAT 321
|||||
GCCTGCCAGAGGAGCTGTGGCAGAGGAGGACGAGCCGTCGGACTGAAT 133
|||||
GACAGAGAAAGCAGGATCCTGGCCCTTCTGTAACCGACTAGTTCGGCTCGC 381
|||||
GACAGAGAAAGCAGGATCCTGGCCCTTCTGTAACCGACTAGTTCGGCTCGC 193
|||||
TGCACTAAAGCCGCGAAACACCGGCTCGAAGAGCGATCGACCCCATATGAA 441
|||||
TGCACTAAAGCCGCGAAACACCGGCTCGAAGAGCGATCGACCCCATATGAA 253
|||||
TCCAGCCTGGAGGAGGAGCGAGGAGGAGTGTGACGGGACAGTGTGGC 501
|||||
TCCAGGACCTGGAGGAGGAGCGAGGAGGAGTGTGACGGGACAGTGTGGC 313
|||||
GGAAGCCAGATCAACAGCTCCAGCCCTCTCGGCTTACAAACCGCCAGATCGGGGAG 561
|||||
GGAAGCCAGATCAACAGCTCCAGCCCTCTCGGCTTACAAACCGCCAGATCGGGGAG 373
|||||
AGTCACCCGGGCTGGGCTTACTACTGTACTGTAGGTGCATTTGTATGAGGGG 621
|||||

Db 374 TTTATAGTCAACCGGCTGGGCTCTACTACCTGTACTGTGAGGTGCACCTTTGATGA(
|||||
QY 622 AAGGCTGTCTACCTGAAGCTGGACTGTGCTGGTGGATGGTGTCTGGCCCTGGGCTG(
|||||
Db 434 AAGGCTGTCTACCTGAAGCTGGACTGTGCTGGTGGATGGTGTCTGGCCCTGGGCTG(
|||||
QY 682 GAGGAATTCTCAGCCACTGGGCGCAGTTCCTCTCGGCGCCCGCAGCTCGGCTCTGCCA(
|||||
Db 494 GAGGAATTCTCAGCCACTGGGCGCAGTTCCTCTCGGCGCCCGCAGCTCGGCTCTGCCA(
|||||
QY 742 TCTGGGCTGTGGCCCTGGGCGCAGG 767
|||||
Db 554 TCTGGGCTGTGGCCCTGGGCGCAGG 579
|||||

RESULT 9
BI871711 731 bp mRNA linear EST 11-
LOCUS 603395825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:54054
DEFINITION mRNA sequence.
ACCESSION BI871711
VERSION BI871711.1 GI:16045386
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 731)
AUTHORS NIH-MGC http://imgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (N
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12034 row: b column: 07
High quality sequence stop: 728.
Location/Qualifiers
1. .731
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405478"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1:
Site 2: SalI; Cloned unidirectionally; oligo-dT p
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technol
Note: this is a NIH_MGC Library."

Query Match 36.9%; Score 506; DB 12; Length 731;
Best Local Similarity 100.0%; Pred. No. 1.8e-246;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps
QY 262 CAGGAGCTGCCAGGAGCTGTGGCAGAGGAGGACGAGCCGTCGGAATTC
|||||
Db 74 CAGGAGCTGCCAGGAGGAGCTGTGGCAGAGGAGGACGAGCCGTCGGAATTC
|||||
QY 322 CCCAGACAGAGAAAGCCAGGATCCTGGCCCTTCTGTAACCGACTAGTTCGGCT
|||||
Db 134 CCCAGACAGAGAAAGCCAGGATCCTGGCCCTTCTGTAACCGACTAGTTCGGCT
|||||
QY 382 AGAAGTGACCTAAAGCCGCGAAACACGGGCTCGAAGAGCGATCGAGCCCATTAI
|||||
```


S.T., Jackson, Y. and Bowers, Y.
ne Pancreas Consortium
shed (2000)
: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
ne Pancreas Consortium
University, Howard Hughes Medical Institute
Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
8

7-495-1812
7-495-8557
dmelton@biochem.harvard.edu
was constructed by Dr. Douglas Melton DNA sequencing by:
ton University Genome Sequencing Center For information on
ng a clone please contact: Juliana Brown
fas.harvard.edu) This sequence now available from the IMAGE
ium, for clone orders contact: info@image.llnl.gov
ality sequence stop: 412.

Location/Qualifiers
1..609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672623"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

35.8%; Score 491; DB 12; Length 609;
rity 100.0%; Pred. No. 7.9e-239;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
TGCGGCTCCCTCGACAGCTCTCTGGGACCCCGTCCCTCTGCCCCACCCCTCA 942
|||||
TGCGGCTCCCTCGACAGCTCTCTGGGACCCCGTCCCTCTGCCCCACCCCTCA 454
|||||
TCCTTGTCTCAGACCTCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTGT 1002
|||||
TCCTTGTCTCAGACCTCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTGT 394
|||||
ATCCACATAAATACAGTATTCACCTCTTATCTTACAACTCCGCCACCCGCCAC 1062
|||||
ATCCACATAAATACAGTATTCACCTCTTATCTTACAACTCCGCCACCCGCCAC 334
|||||
ACCTCAGTAGTCCCAATCCCTGACCCCTTGGGCCCCCAGTAGTCTGACCTCC 1122
|||||
ACCTCAGTAGTCCCAATCCCTGACCCCTTGGGCCCCCAGTAGTCTGACCTCC 274
|||||
GGCCACAGACCCCCAGGAGGCTGTGTTACCTGTACTCTGTGGGCAAGGATGGGTC 1182
|||||
GGCCACAGACCCCCAGGAGGCTGTGTTACCTGTACTCTGTGGGCAAGGATGGGTC 214
|||||
GACCCCACTTCAGGACCTAAGAGGGGCTGGACCTTGGCGGCGAGGAGCAAGAGA 1242
|||||
GACCCCACTTCAGGACCTAAGAGGGGCTGGACCTTGGCGGCGAGGAGCAAGAGA 154
|||||
CCTAGGCCAGGAGTTCCTCAATGTGAGGGGCGAGAAACAAGCAAGCTCTCTCCCT 1302

Db 153 CTGGGCTAGGCGAGGAGTCCCAATCTGAGGGCGAGAAACAAGCAAGCTCCTC
|||
Qy 1303 TGAGATTCCTCTGGATTTTAAACAGATATTTTATTTATTTATTTGACAA
|||
Db 93 TGAGATTCCTCTGGATTTTAAACAGATATTTTATTTATTTATTTGACAA
|||
Qy 1363 TTGATAAATGG 1373
|||
Db 33 TTGATAAATGG 23
|||

RESULT 12
CF126539 666 bp mRNA linear EST 05-
LOCUS
DEFINITION UI-HF-ET0-avw-m-22-0-UL.r1 NIH MGC_214 Homo sapiens cDNA cl
IMAGE:30559317 5', mRNA sequence.
ACCESSION CF126539
VERSION CF126539.1 GI:33203878
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 666)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 97044477
MEDLINE 8889548
PUBMED
COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
cDNA Library preparation: Dr. M. Bento Soares, University
cDNA Library Arrayed by: Dr. M. Bento Soares, University
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pyX-5.

FEATURES
Location/Qualifiers
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30559317"
/tissue_type="Chondrosarcoma Lung Metastasis cell
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH MGC_214"
/note="Organ: Lung; Vector: pyX-Asc; Site 1: EcoR
Site 2: Not I; The library was constructed accordi
Bonaldo, Lennon and Soares, Genome Research, 6:791
1996. Denatured RNA was size fractionated on a 1%
gel. First strand cDNA synthesis was primed with c
primer containing a Not I site. Double strand cDNA
size selected according to mRNA size fraction, lig
with EcoR I adaptor, digested with Not I and then
directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the po
is GATAAGGCCA. Tissue was provided by Mary Hendrix

ORIGIN

Query Match 33.9%; Score 465; DB 14; Length 666;
Best Local Similarity 99.6%; Pred. No. 1.5e-225;
Matches 565; Conservative 0; Mismatches 2; Indels 0; Ga
Qy 303 GGACCCGTGGAACTGATCCCGACAGAGAGGAGGAGGATCTGCGCTTCCT
|||||

CCGTCGGAAGTAAATCCACAGAGAAAGCCAGGATCCTGGCCCTTCTCGAA 71
 CTAGTTCGGCTCGCAGAGTGCACCTAAAGCCCGGAAACACACGGGTGGAAGAGC 422
 CTAGTTCGGCTCGAAGAGTGCACCTAAAGCCCGGAAACACACGGGTGGAAGAGC 131
 GCAGCCCATTAATGAAGTTTCATCCAGACTCGACAGGACGAGCGCAGGAGGTGT 482
 GCAGCCCATTAATGAAGTTTCATCCAGACTCGACAGGACGAGCGCAGGAGGTGT 191
 GGGCAGTGAAGTGGTGGGAGGAGCCAGATCAACAGCTCCAGCCCTTCGGCTA 542
 GGGCAGTGAAGTGGTGGGAGGAGCCAGATCAACAGCTCCAGCCCTTCGGCTA 251
 GGCAGATCGGGAGTTTATAGTCAACCCGGCTGGCTCTACTACTGTACTGTCA 602
 GGCAGATCGGGAGTTTATAGTCAACCCGGCTGGCTCTACTACTGTACTGTCA 311
 GACTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTGTCTGGTGGTGT 371
 GGCCTCGCTCGCTGGGAGGATCTCAGCCACTGGGCGAGTTCCTCGGCGCCA 722
 GGCCTCGCTCGCTGGGAGGATCTCAGCCACTGGGCGAGTTCCTCGGCGCCA 431
 GGCCTCGCTCGCTGGGAGGATCTCAGCCACTGGGCGAGTTCCTCGGCGCCA 782
 GGCCTCGCTCGCTGGGAGGATCTCAGCCACTGGGCGAGTTCCTCGGCGCCA 491
 ACCCTCCCTCGGCGGATCTCAAGGCTCGGCGGCTTCCTCACTACTCGGACTCT 842
 ACCCTCCCTCGGCGGATCTCAAGGCTCGGCGGCTTCCTCACTACTCGGACTCT 551
 GTTCACTGAGGGGCGCTGGTCTC 869
 GTTCACTGAGGGGCGCTGGTCTC 578

15.1 HR85 islet Homo sapiens cDNA clone IMAGE:6128001 3',
 sequence.
 15.1 GI:24203667
 sapiens (human)
 sapiens
 Meta; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,
 ka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 T., L., Marra, M., Pape, D., Wylie, T., Martin, J., Bliscain, A.,
 t, A., Theising, B., Ratter, E., Ronko, I., Bennett, J.,
 as, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,
 ms, T., Jackson, Y., and Bowers, Y.
 ine Pancreas Consortium
 ished (2000)
 t: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 ine Pancreas Consortium
 d University, Howard Hughes Medical Institute
 of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 38
 17-495-1812
 17-495-8557
 dmelton@biohp.harvard.edu
 y was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 ington University Genome Sequencing Center For information on
 ing a clone please contact: Dr. Hiroshi Inoue
 e@im.wustl.edu)
 1mer: -400P from Gibco

High quality sequence stop: 451.
 Location/Qualifiers
 1. 474
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6128001"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-)
 Noti; Site 2: XhoI; cDNA made by oligo-dT priming
 Size selected on agarose gel. Average insert size
 XhoI site was destroyed after directional cloning
 Amplified once. Contact information: Hiroshi Inoue
 Metabolism Div. (Alan Permut Lab), Washington Ur
 School of Medicine, Box 8127, 560 South Euclid Av
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu,
 314-362-1916, Fax: 314-747-2692."

ORIGIN
 Query Match 33.6%; Score 462; DB 13; Length 474;
 Best Local Similarity 100.0%; Pred No. 4.8e-224;
 Matches 462; Conservative 0; Mismatches 0; Indels 0; G

QY 912 TGGCACCCTGGTCCCTCTGCCCCACCTCAGCCGCTCTTTGCTCCAGACCTGCC 474
 Db 474 TGGCACCCTGGTCCCTCTGCCCCACCTCAGCCGCTCTTTGCTCCAGACCTGCC 474
 QY 972 CTCTAGAGGCTGCTGGGCTGTTTACGTGTTTCCATCCACATAATACAGTAT 414
 Db 414 CTCTAGAGGCTGCTGGGCTGTTTACGTGTTTCCATCCACATAATACAGTAT 414
 QY 1032 ACTCTATCTTACAACTCCCCACCGCCACTCTCCACTCACTAGCTCCCAATC 354
 Db 354 ACTCTATCTTACAACTCCCCACCGCCACTCTCCACTCACTAGCTCCCAATC 354
 QY 1092 ACCCTTTGAGGCCCCAGTGTCTGACTCCCCCTGGCCACAGACCCCAAGGGC 294
 Db 294 ACCCTTTGAGGCCCCAGTGTCTGACTCCCCCTGGCCACAGACCCCAAGGGC 294
 QY 1152 GTTCACTGTACTCTGGGCAAGGATGGTCCAGAGACCCCACTTCAGGCACTAT 234
 Db 234 GTTCACTGTACTCTGGGCAAGGATGGTCCAGAGACCCCACTTCAGGCACTAT 234
 QY 1212 GCTGGACCTGGCGCAGAGCAAGCAAGAGACTGGGCTAGGCCAGGAGTTCCCAAT 174
 Db 174 GCTGGACCTGGCGCAGAGCAAGCAAGAGACTGGGCTAGGCCAGGAGTTCCCAAT 174
 QY 1272 GAGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAAATTCCTGTGGATTTTAA 114
 Db 114 GAGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAAATTCCTGTGGATTTTAA 114
 QY 1332 ATATTATTTTATTATTATTATTGTGACAAATGTTGATAATGG 1373
 Db 54 ATATTATTTTATTATTATTATTGTGACAAATGTTGATAATGG 13

RESULT 14
 CB529199/c
 LOCUS
 DEFINITION
 UI-H-FT2-bj-f-k-03-0-UI.s1 NCI CGAP_FT2 Homo sapiens cDNA c
 UI-H-FT2-bj-f-k-03-0-UI 3', mRNA sequence.
 ACCESSION
 CB529199
 VERSION
 CB529199.1 GI:29389647
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 569)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS
 TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (

me Index
shed (1997)
Robert Strausberg, Ph.D.
cgapbs-remail.nih.gov
Procurement: Dr. Gary W. Hunninghake, U of I
Library preparation: Dr. M. Bento Soares, University of Iowa
Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Sequencing by: Dr. M. Bento Soares, University of Iowa
Distribution: Distribution information can be found at
genome.uiowa.edu/distribution/cgap.html
Following repetitive elements were found in this cDNA
1-82, >AT rich/Low complexity (matched complement)
er: M13 FORWARD
as.

Location/Qualifiers
1..569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT2-bj-f-k-03-0-UI"
/tissue_type="Aveolar Macrophage"
/dev stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FT2"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP FT2 is a subcloned cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
subtracted according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. The tissue was provided by Dr.
Gary W. Hunninghake of the University of Iowa.
TAG TISSUE=Human Lung Aveolar Macrophage
TAG LIB=UI-H-FT2
TAG_SEQ=GGCATGCGG"

32.7%; Score 449; DB 14; Length 569;
rity 99.8%; Pred. No. 2.1e-217;
nservative 0; Mismatches 1; Indels 0; Gaps 0;

3TCCAGGCTGCGGCTCCCTCGACAGCTCTCTGGGACCGGCTCCCTCTGCG 933
3TCCAGGCTGCGGCTCCCTCGACAGCTCTCTGGGACCGGCTCCCTCTGCG 468
CTGAGCGCTTTGCTCGAGCTGCGGCTCCCTCTAGAGGCTGCTGGGCTG 993
CTGAGCGCTTTGCTCGAGCTGCGGCTCCCTCTAGAGGCTGCTGGGCTG 408
GTGTTTTCATCCACATAATACAGTATTCCCACTCTTATCTTACAACTCCCCC 1053
GTGTTTTCATCCACATAATACAGTATTCCCACTCTTATCTTACAACTCCCCC 348
CCACTCTCCAGCTCAGTCTCCCAATCCCTGACCCCTTTGAGCCGCCAGTAGT 1113
CCACTCTCCAGCTCAGTCTCCCAATCCCTGACCCCTTTGAGCCGCCAGTAGT 288
CTCCCCCTGSCACACAGACCCCGAGGCTGTTGTTCACTGTACTCTGTGGGCAA 1173
CTCCCCCTGSCACACAGACCCCGAGGCTGTTGTTCACTGTACTCTGTGGGCAA 228
GGTCCAGAACACCCCACTTCAGGCACTTAAGAGGGGCTGGACCTTGGCGGAGGAG 1233
GGTCCACAGACCCCACTTCAGGCACTTAAGAGGGGCTGGACCTTGGCGGAGGAG 168
GAGACTGGGCTAGGCGAGGAGTTCCCAATGTGAGGGGGGAGAAACAAGCAAG 1293
GAGACTGGGCTAGGCGAGGAGTTCCCAATGTGAGGGGGGAGAAACAAGCAAG 108
CCCTTGAGATTCCTCTGGATTTTAAACAGATATATTTTATTATTATTGT 1353
CCCTTGAGATTCCTCTGGATTTTAAACAGATATATTTTATTATTATTGT 48

QY 1354 GACAAATGTTGATAAATGG 1373
Db 47 GACAAATGTTGATAAATGG 28
RESULT 15
BQ707185
LOCUS
DEFINITION BQ707185 948 bp mRNA linear EST 16-
AGENCOURT 8353983 NIH_MGC_113 Homo sapiens cDNA clone IMAGE
5', mRNA sequence.
ACCESSION BQ707185
VERSION BQ707185.1 GI:21846084
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleo
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 948)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (M
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLCM2466 row: n column: 17
High quality sequence start: 24
High quality sequence stop: 550.
Location/Qualifiers
1..948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6278608"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI;
EcoRI; cDNA made by oligo-dT priming. Directional
into EcoRI/XhoI sites using the following 5' adapt
GGCAGCAG(G). Library constructed by Ling Hong in
laboratory of Gerald M. Rubin (University of Cali
Berkeley) using ZAP-cDNA synthesis kit (Stratagene
Superscript II RT (Life Technologies)). Note: this
NIH_MGC Library."

ORIGIN

Query Match 31.8%; Score 437; DB 13; Length 948;
Best Local Similarity 99.8%; Pred. No. 3.2e-211;
Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps
QY 475 GCAGGTGTGGACGGGACAGTGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC
Db 167 GCAGGTGTGGACGGGACAGTGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC
QY 535 CTGCGCTCAACCGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTCTACTAC
Db 227 CTGCGCTCAACCGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTCTACTAC
QY 595 TACTGTCTGAGTGCATTTGATGAGGGGAGGCTGTCTACCTGAACTGACCTTGCTG
Db 287 TACTGTCTGAGTGCATTTGATGAGGGGAGGCTGTCTACCTGAACTGACCTTGCTG
QY 655 GATGTGTGTGCTGGCCCTGCGCTGCTGAGGAAATTTCTCAGCCACTGCGGCCAGTTCC
Db 347 GATGTGTGTGCTGGCCCTGCGCTGCTGAGGAAATTTCTCAGCCACTGCGGCCAGTTCC
QY 715 GGGCCCCAGCTCGCCCTCTGCCAGGTGTCTGGGCTGTGGGCTCGGCTCGGCGGCTCC

CCAGCTCCGCTCTGCGAGGTGCTGGGCTGTGGCCCTTCGGCCAGGGTCTCTCC 466
 GGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCACTACTTC 834
 GGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCACTACTTC 526
 TCTTCCAGGTTCACTGAGGGCCCTGTGTCTCCACAGTGTCTCCAGGCTGCCGGC 894
 TCTTCCAGGTTCACTGAGGGCCCTGTGTCTCCCGCAGTGTCTCCAGGCTGCCGGC 586
 CTCACAGCTCTCTGGGACCCGCTGCCCTCTGCCACACCTCACTGAGCGCTCTTTC 954
 CTCACAGCTCTCTGGGACCCGCTGCCCTCTGCCACACCTCACTGAGCGCTCTTTC 646
 GACC 962
 GACC 654

31 940 bp mRNA linear EST 16-AUG-2002
 URT 8682031 Lupski_sciatic_nerve Homo sapiens cDNA clone
 6197488 5', mRNA sequence.
 31
 31.1 GI:22276239
 sapiens (human)

apiens
 ata; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ses 1 to 940)
 C http://mgc.nci.nih.gov/
 al Institutes of Health, Mammalian Gene Collection (MGC)
 ished (1999)
 t: Robert Strausberg, Ph.D.
 cgaps-remail.nih.gov
 Procurement: Dr. James R. Lupski
 Library Preparation: Life Technologies, Inc.
 Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 sequencing by: Agencourt Bioscience Corporation
 distribution: MGC clone distribution information can be
 through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov
 LLAM13607 row: j column: 17
 uality sequence stop: 453.
 Location/Qualifiers

1..940
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6197488"
 /sex="male"
 /tissue_type="sciatic nerve"
 /dev_stage="adult, 70 Yr"
 /lab_host="DH10B"
 /clone_lib="Lupski sciatic nerve"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
 Not1, Site 2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCTCAGGCTCCG-3' and
 5'-GACTAGTCTAGATCGAGCGCGCCCT(15)-3'. Size selected >
 1 kb for average insert length 1.87 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

31.8%; Score 436; DB 13; Length 940;
 arity 100.0%; Pred.No.1e-210;
 onservative 0; Mismatches 0; Indels 0; Gaps 0;

301 CAGGACCCGTCGGAATGCCAGACAGAAAGCCAGGATCCTGCGCCTTT
 1 CAGGACCCGTCGGAATGCCAGACAGAAAGCCAGGATCCTGCGCCTTT
 361 AACGACTAGTTGGCCTCGCAGAGTGACCTAAAGCCGCGGAAACACGGGCTCG
 61 AACGACTAGTTGGCCTCGCAGAGTGACCTAAAGCCGCGGAAACACGGGCTCG
 421 GCATCCGACGCCATTATGAAGTTTCATCCAGCTCGACAGGACGAGGCGCAGGC
 121 GCATCCGACGCCATTATGAAGTTTCATCCAGCTCGACAGGACGAGGCGCAGGC
 481 GTGACCGGACAGTGTGCTGGAGGAAAGCCAGAAATCAACAGCTCCAGCCCTCT
 181 GTGACCGGACAGTGTGCTGGAGGAAAGCCAGAAATCAACAGCTCCAGCCCTCT
 541 TACAACCGCAGATCGGGGAGTTTATAGTACCCGGGCTGGCTCTACTACTCTGA
 241 TACAACCGCAGATCGGGGAGTTTATAGTACCCGGGCTGGCTCTACTACTCTGA
 601 CAGGTGCACCTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTCTGCTGGA
 301 CAGGTGCACCTTTGATGAGGGAAGGCTGTCTACCTGAAGCTGGACTTCTGCTGGA
 661 GTGCTGGCCCTGCGCTGCGTGGAGGAAATTCACGCCACTCGGCGCAGTTCCCTCGG
 361 GTGCTGGCCCTGCGCTGCGTGGAGGAAATTCACGCCACTCGGCGCAGTTCCCTCGG
 721 CAGCTCCGCTCTGCC 736
 421 CAGCTCCGCTCTGCC 436

RESULT 17
 CF126932 597 bp mRNA linear EST 05
 LOCUS
 DEFINITION
 UI-HF-ETO-avx-k-19-0-UI.r1 NIH_MGC_214 Homo sapiens cDNA c
 IMAGE:30563490 5', mRNA sequence.

CF126932
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 697)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitat
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Mary Hendrix
 cDNA Library Preparation: Dr. M. Bento Soares, University
 cDNA Library Arrayed by: Dr. M. Bento Soares, University
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found
 http://genome.uiowa.edu/distribution/humanfl.html
 The following repetitive elements were found in this cDN
 sequence: 37-143, >GC-rich#Low_complexity
 Seq primer: pYX-5.

FEATURES
 source
 1..697
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

/clone="IMAGE:30563490"
 /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_MGC_214"
 /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GATAGGCCA. Tissue was provided by Mary Hendrix."

31.0%; Score 425; DB 14; Length 697;
 rity 100.0%; Pred. No. 4e-205;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 GCCTGCCAGGAGGAGCTGTGGCAGAGGAGCAGGACCCGTCGGAACCTGAAT 321
 |||||
 GCCTGCCAGGAGGAGCTGTGGCAGAGGAGCAGGACCCGTCGGAACCTGAAT 268
 |||||
 GACAGAAGAAGCCAGGATCCTGCGCCTTTCTGAAACCGACTAGTTCGGCCTCGC 381
 |||||
 GACAGAAGAAGCCAGGATCCTGCGCCTTTCTGAAACCGACTAGTTCGGCCTCGC 328
 |||||
 TGCACCTAAAGCCGGAAACACAGGCTCGAAGAGCGATCGACCCCATTTATGAA 441
 |||||
 TGCACCTAAAGCCGGAAACACAGGCTCGAAGAGCGATCGACCCCATTTATGAA 388
 |||||
 TCACGACCTGGACAGAGGAGCGAGCGAGCGAGCGAGTGTGACCGGACATGAGTGC 448
 |||||
 GGAAAGCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATCGGGAG 508
 |||||
 AGTCACCGGGCTGGCTCTACTACCTGTACTGTACCTGTACTGTACCTGTACTGT 621
 |||||
 AGTCACCGGGCTGGCTCTACTACCTGTACTGTACCTGTACTGTACCTGTACTGT 568
 |||||
 TGCTACTGAAGCTGAGCTGTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCGCTG 681
 |||||
 TGCTACTGAAGCTGAGCTGTGCTGGTGGATGGTGTGCTGGCCCTGCGCTGCGCTG 628
 |||||

686

633

6 834 bp mRNA linear EST 25-SEP-2001
 66F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
 quence.

6.1 GI:15758344

piens (human)
 piens
 ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 es 1 to 834)
 http://img.ncbi.nih.gov/
 l Institutes of Health, Mammalian Gene Collection (MGC)
 shed (1999)
 : Robert Strausberg, Ph.D.
 cgabps-r@mail.nih.gov
 Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information car
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

High quality sequence stop: 772.
 Plate: LLAM11517 row: c column: 18

FEATURES

Location/Qualifiers
 1..834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5206217"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_122"
 /note="Organ: pooled lung and spleen; Vector: pCMV
 Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week fem
 spleen, and 20-22 week male spleens. Library is c
 primed and directionally cloned (EcoRV site is des
 upon cloning). Average insert size 1.4 kb. inser
 range 1-3 kb. Library is normalized and enriched f
 full-length clones and was constructed by C. Gruk
 (Invitrogen). Research Genetics tracking code 02
 this is a NIH_MGC Library."

ORIGIN

Query Match 30.3%; Score 416; DB 12; Length 834;
 Best Local Similarity 99.7%; Pred. No. 1.6e-200;
 Matches 656; Conservative 0; Mismatches 0; Indels 2; Gaps
 QY 272 CCAGGAGGAGCTGTGGCAGAGGAGGAGCAGGACCCGTCGGAACCTGAATCCCCAGF
 Db 1 CCAGGAGGAGCTGTGGCAGAGGAGGAGCAGGACCCGTCGGAACCTGAATCCCCAGF
 QY 332 AAGAAAGCCAGGATCCTGCGCCTTTCTGAAACCGACTAGTTCGGCCTCGCAGAAATGTC
 Db 61 AAGAAAGCCAGGATCCTGCGCCTTTCTGAAACCGACTAGTTCGGCCTCGCAGAAATGTC
 QY 392 CTAAGGCGGGAACACAGGGCTCGAAGAGCGATCGAGCCCATTTATGAAGTTCATC
 Db 121 CTAAGGCGGGAACACAGGGCTCGAAGAGCGATCGAGCCCATTTATGAAGTTCATC
 QY 452 GACC-TGGACAGGACCGAGCGCAGGAGTGTGGACGGGACAGTCAAGTGGCTGGGAC
 Db 181 GACCGTGACAGGAGGAGCGCAGGAGTGTGGACGGGACAGTCAAGTGGCTGGGAC
 QY 511 GCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATCGGGAGTTTATP
 Db 241 GCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATCGGGAGTTTATP
 QY 571 ACCCGGGCTGGGCTCTACTACTGTGATGTGTGCTGGCCCTCGCTGCTGCTGAGGAA
 Db 301 ACCCGGGCTGGGCTCTACTACTGTGATGTGTGCTGGCCCTCGCTGCTGCTGAGGAA
 QY 631 TACCTGAAGCTGGACTTCTGCTGGTGGATGTGTGCTGGCCCTCGCTGCTGCTGAGGAA
 Db 361 TACCTGAAGCTGGACTTCTGCTGGTGGATGTGTGCTGGCCCTCGCTGCTGCTGAGGAA
 QY 691 TCAGCACTGCGGCGCAGTTCCCTCGGGCCCGCAGCTCCGCTTCGCAAGGTGTCTGGC
 Db 421 TCAGCACTGCGGCGCAGTTCCCTCGGGCCCGCAGCTCCGCTTCGCAAGGTGTCTGGC
 QY 751 TTGGCCCTCGGCGCAGGGTCTCCCTCGGATCGGATCGGATCGGATCGGATCGGATCGG
 Db 481 TTGGCCCTCGGCGCA-GGTCTCTCTCGGATCGGATCGGATCGGATCGGATCGGATCGG
 QY 811 GCTGCCCCCTTCTCCTCACCTACTTCTGAGTCTTCCAGGTTCACTGAGGGGCCCTGTGTC
 Db 540 GCTGCCCCCTTCTCCTCACCTACTTCTGAGTCTTCCAGGTTCACTGAGGGGCCCTGTGTC
 QY 871 CCACAGTGTGTCAGGCTGCGGGTCCCTCGACAGCTCTCTGCGGACCCCGGTCCCG


```
|||||
AGTCGCCAGGCTGCCGCTCCCTCGACAGCTCTCTGGGCAACCCGCTCCCT 657

96 6.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2096962 3',
sequence.
96 96
96.1 GI:4268727
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap,
al Cancer Institute / National Institute of Neurological
ers and Stroke, Brain Tumor Genome Anatomy Project
BTGAP), Tumor Gene Index
ished (1996)
t: Robert Strausberg, Ph.D.
cgapbs-remail.nih.gov
Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
o, Ph.D.
Library Preparation: M. Bento Soares, Ph.D., M. Fatima
o, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
sequencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
o.llnl.gov/bbrp/image/image.html
Length: 1728 Std Error: 0.00
imer: -40UP from Gibco
quality sequence stop: 410.
Location/Qualifiers
1..413
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2096962"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP_Brn23"
/notes="Organ: Brain; Vector: pTT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCATCTGAAGTGGGAGCGCGCATATCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pTT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
28.7%; Score 394; DB 9; Length 413;
arity 100.0%; Pred. No. 2.4e-189;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
|CTGGGCTGTTACGCTGTTTCCATCCACATATACATATCCACCTTTAT 1039
|CCTGGGCTGTTTACGCTGTTTCCATCCACATATATACATATCCACCTTTAT 354
|CAACTCCCCACCGCCCACTCTCCACCTCAGTCTCCCAATCCCTGACCCCTTG 1099
|CTGGGCTGTTTACGCTGTTTCCATCCACATATACATATCCACCTTTAT 1039
|CAACTCCCCACCGCCCACTCTCCACCTCAGTCTCCCAATCCCTGACCCCTTG 294
|CCCCAGTATCTGACTCCCCCTGGGCCACAGACCCCGAGGCGATGTGTACATG 1159
|CCCCAGTATCTGACTCCCCCTGGGCCACAGACCCCGAGGCGATGTGTACATG 234
|CTGTGGGAGGATGGTCCAGAACCCCACTTCAGGCACTAAGAGGGGCTGGAC 1219
```

```
|||||
233 TACTCTGTGGGCAAGGATGGTCCAGAAAGCCCACTTCAGGCACCTAAGAGGGGCT
|||||
1220 CTGGCGGCGAGGAGCCAAAGAGAGCTGGGCGCTAGGCCAGGAGTTCCCAAAATGTGAGG
|||||
173 CTGGCGGCGAGGAGCCAAAGAGAGCTGGGCGCTAGGCCAGGAGTTCCCAAAATGTGAGG
|||||
1280 AGAAACAAGACAGAGCTCTCTCCCTTGAGAAATCCCTGTGGATTTTAAACACAGATAT
|||||
113 AGAAACAAGACAGAGCTCTCTCCCTTGAGAAATCCCTGTGGATTTTAAACACAGATAT
|||||
1340 TTATTATTATTGTGACAAATGTTGATAATGG 1373
53 TTATTATTATTGTGACAAATGTTGATAATGG 20

RESULT 20
BM971606/c
LOCUS
DEFINITION
UI-CF-EC1-abl-p-06-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clc
UI-CF-EC1-abl-p-06-0-UI 3', mRNA sequence.
BM971606
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 568)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitat
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USF
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University
cDNA Library Arrayed by: Dr. M. Bento Soares, University
DNA Sequencing by: Dr. M. Bento Soares, University of Iow
Clone Distribution: Researchers may obtain clones from Re
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDN
sequence: 1-82, >AT-rich#low_complexity (matched complim
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1..568
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-abl-p-06-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage re
/clone_lib="UI-CF-EC1"
/notes="Organ: Lung; Vector: pTT73-Pac (Pharmacia)
modified polylinker; Site 1: Eco RI; Site 2: Not
UI-CF-EC1 is a normalized cDNA library containing
following tissue(s): Normal lung from adult and
day 64, day 87, week 19 and week 42. The library
constructed according to Bonaldo, Lennon and Soa
Genome Research, 6:791-806, 1996. First strand cl
synthesis was primed with an oligo-dT primer cont
Not I site. Double stranded cDNA was ligated to
adaptor, digested with Not I, and cloned directic
```

into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (GT)₁₈ tail. The sequence tag for this library is AAGTGGTTAC.
TAG TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371 and 380-383
TAG LIB=UI-CF-EC1
TAG_SEQ=AAGTGGTTAC"

28.7%; Score 394; DB 12; Length 568;
rity 100.0%; Pred. No. 2.5e-189;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
CTGGGCGTTCACGTGTTTCCATCCACATAAATACAGTATCCCACTCTTAT 1039
CTGGGCGTTCACGTGTTTCCATCCACATAAATACAGTATCCCACTCTTAT 362
AACTCCCCCAGCGCCACTCTCCACCTCAGTCTCCCAATCCCTGACCCCTTG 1099
AACTCCCCCAGCGCCACTCTCCACCTCAGTCTCCCAATCCCTGACCCCTTG 302
CCAGTATCTCGACTCCCTCCCTGGCCACAGACCCCGAGGCATTGTTCACGTG 1159
CCAGTATCTCGACTCCCTCCCTGGCCACAGACCCCGAGGCATTGTTCACGTG 242
TGTGGCAAGGATGGTCCAGAACCCCTTCCAGGACCTAAGAGGGCTGGAC 1219
TGTGGCAAGGATGGTCCAGAACCCCTTCCAGGACCTAAGAGGGCTGGAC 182
GGCAGGAGCCAAAGAGACTGGGCTTAGCCAGGAGTTCCCAATGTGAGGGCG 1279
GGCAGGAGCCAAAGAGACTGGGCTTAGCCAGGAGTTCCCAATGTGAGGGCG 122
CAAGCAAGCTCCCTCCCTTGAGAAATTCCTGTGAAATTTTAAACAGATATTAT 1339
CAAGCAAGCTCCCTCCCTTGAGAAATTCCTGTGAAATTTTAAACAGATATTAT 62
TATTATTGTGACAAATGTCATAAATGG 1373
TATTATTGTGACAAATGTCATAAATGG 28

4 569 bp mRNA linear EST 23-SEP-2002
1-bdu-c-24-0-UI.s1 NCI CGAP FE1 Homo sapiens cDNA clone
1-bdu-c-24-0-UI 3', mRNA sequence.
4
4.1 GI:23298519
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 569)
P http://www.ncbi.nlm.nih.gov/ncicgap.
1 Cancer Institute, Cancer Genome Anatomy Project (CGAP),
ene Index
shed (1997)
: Robert Strausberg, Ph.D.
cgapsb-r@mail.nih.gov
Procurement: James Martin
Library preparation: Dr. M. Bento Soares, University of Iowa
Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Sequencing by: Dr. M. Bento Soares, University of Iowa
Distribution: Clone distribution information can be obtained
from M. Bento Soares, bentso-soares@uiowa.edu
allowing repetitive elements were found in this cDNA
re: 1-82, >AT-rich/low complexity (matched complement)
mer: M13 FORWARD
es.

FEATURES

Location/Qualifiers
1..569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FE1-bdu-c-24-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FE1"
/note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ec Site 2: Not I; NCI CGAP FE1 is a normalized cDNA 1 derived from a pool of mRNA obtained from 3 cell 1 from grade II chondrosarcoma tissues. The library constructed according to Bonaldo, Lennon and Soares Genome Research, 8:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer conta Not I site. Double stranded cDNA was ligated to a adaptor, digested with Not I, and cloned director into pT73-Pac vector. The oligonucleotide used to the synthesis of first-strand cDNA contains a libi sequence that is located between the Not I site ar (dT)₁₈ tail. The sequence tag for this library is CGTACGGAC. The cell lines were provided by Dr Jan Martin from the University of Iowa.
TAG TISSUE=Human grade 2 chondrosarcoma cell line
TAG LIB=UI-H-FE1
TAG_SEQ=CGTACGGAC"

ORIGIN

Query Match . 28.6%; Score 393; DB 13; Length 569;
Best Local Similarity 99.8%; Pred. No. 8.2e-189;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps
QY 930 TGCCCCACCTCAGCGCTCTTTGTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCC
DB 471 TGCCCCACCTCAGCGCTCTTTGTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCC
QY 990 CCGTGTTCAGTGTCTTCCATCCACATAAATACAGTATCCCACTCTTATCTTACAT
DB 411 CCGTGTTCAGTGTCTTCCATCCACATAAATACAGTATCCCACTCTTATCTTACAT
QY 1050 CCCCACCGCCCACTCTCCACCTCAGTCTCCCCCAATCCCTGACCCCTTTGAGGCGCC
DB 351 CCCCACCGCCCACTCTCCACCTCAGTCTCCCCCAATCCCTGACCCCTTTGAGGCGCC
QY 1110 TGAATCTGACTCCCCCTGGCCACAGACCCCGAGGCATTGTTCACGTACTCT
DB 291 TGAATCTGACTCCCCCTGGCCACAGACCCCGAGGCATTGTTCACGTACTCT
QY 1170 GCAAGGATGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGCTGGACCTGGCG
DB 231 GCAAGGATGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGCTGGACCTGGCG
QY 1230 GAAGCCAAAGAGACTGGGCTTAGGCCAGGAGTTCCCAATGTGAGGGCGGAGAAAC
DB 171 GAAGCCAAAGAGACTGGGCTTAGGCCAGGAGTTCCCAATGTGAGGGCGGAGAAAC
QY 1290 CAACTCTCCCTTCAGAAATTCCTGTGAAATTTTAAACAGATATTATTTTATT
DB 111 CAACTCTCCCTTCAGAAATTCCTGTGAAATTTTAAACAGATATTATTTTATT
QY 1350 TTGTGACAAATGTTGATAAATGG 1373
DB 51 TTGTGACAAATGTTGATAAATGG 28

RESULT 22
BI824443
LOCUS
DEFINITION
60308693F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179:
mRNA sequence.

/clone_lib="NIH_MGC_102"
 Note: Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed
 by Ling Hong in the laboratory of Gerald M. Rubin
 (University of California, Berkeley) using ZAP-cDNA
 synthesis kit (Stratagene) and Superscript II RT (Life
 Technologies). Note: this is a NIH_MGC Library."

25.4%; Score 349; DB 13; Length 951;
 arity 100.0%; Pred. No. 2.6e-166;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CGGCTCGCAGAGTGCACCTAAAGCGCGAAACACGGCTCGAAGCGATCGC 428
 CGGCTCGCAGAGTGCACCTAAAGCGCGAAACACGGCTCGAAGCGATCGC 60
 CATTATGAAGTTTCATCCACGACCTGGACAGGACGCGCAGGCGAGTGTGGACGG 488
 CATTATGAAGTTTCATCCACGACCTGGACAGGACGCGCAGGCGAGTGTGGACGG 120
 GTGAGTGGCTGGGAGGAGCGAGCAATCAACAGCTCCAGCCCTCTCGCTACAACCG 548
 GTGAGTGGCTGGGAGGAGCGAGCAATCAACAGCTCCAGCCCTCTCGCTACAACCG 180
 ATCCGGGAGTTTATAGTCAACCGGGCTGGCTCTACTACCTGTACTGTTCAGGTGCA 608
 ATCCGGGAGTTTATAGTCAACCGGGCTGGCTCTACTACCTGTACTGTTCAGGTGCA 240
 GATGAGGGGAAGGCTGTCTACTGAAGCTGGACTTGTGTGGTGTGTGTGTGTGTGTGT 668
 GATGAGGGGAAGGCTGTCTACTGAAGCTGGACTTGTGTGGTGTGTGTGTGTGTGTGT 300
 CGCTGCTCGGAGGAATTCAGCCACTCGCGCCAGTTCCTTCGGG 717
 CGCTGCTCGGAGGAATTCAGCCACTCGCGCCAGTTCCTTCGGG 349

16 377 bp mRNA linear EST 15-FEB-2002
 5.Y1 Human insulinoma Homo sapiens cDNA 5', mRNA sequence.
 16
 16.1 GI:18680159

sapiens (human)
 Meta; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (see 1 to 377)
 L.D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,
 ka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 L., Marra, M., Pape, D., Wylie, T., Martin, J., Bliscain, A.,
 T.A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
 as, M., Gibbons, M., McCann, R., Cole, R., Teagareishvili, R.,
 ms, T., Jackson, Y. and Bowers, Y.
 ine Pancreas Consortium
 ished (2000)
 ESTs: ih15605.x1
 it: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 ine Pancreas Consortium
 d University, Howard Hughes Medical Institute
 of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 [38
 317-495-1812
 317-495-8557
 : dmelton@biohpc.harvard.edu
 cy was constructed by Dr. J. J. Ferrer in vivo mass-excised to
 script SK- by Dr. H. Inoue DNA sequencing by: Washington
 :sity Genome Sequencing Center for information on obtaining a
 please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco.
 Location/Qualifiers
 1. 377
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="insulinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Human insulinoma"
 /note="Organ: pancreas; Vector: pBluescript SK-;
 XhoI; Site 2: EcoRI; Constructed with lambda ZAPI
 pBluescript SK- by Dr. H. Inoue following the Wa
 University protocol
 (http://genome.wustl.edu/est/lambda_protocol.shtm
 please contact Hiroshi Inoue, MD/PhD for further
 information on this library (Metabolism Division
 Laboratory, Washington University School of Medi
 8127, 660 S Euclid Ave, St. Louis, MO 63110). NC
 is a Washington University Pancreas EST project 1

ORIGIN

Query Match 24.8%; Score 341; DB 12; Length 377;
 Best Local Similarity 100.0%; Pred. No. 2.6e-162; Indels 0; C
 Matches 341; Conservative 0; Mismatches 0;
 QY 1033 CTCCTATCTTACAACTCCCGCCAGCTCTCCACCTCACTAGCTCCCAATCC
 DB 8 CTCCTATCTTACAACTCCCGCCAGCTCTCCACCTCACTAGCTCCCAATCC
 QY 1093 CCCTTTGAGGCCCCAGTGATCTCGACTCCCGCTGGCCAGACCCCGAGGCA
 DB 68 CCCTTTGAGGCCCCAGTGATCTCGACTCCCGCTGGCCAGACCCCGAGGCA
 QY 1153 TTCACTGTACTCTGTGGCAAGGATGGTCCAGAAAGACCCCACTTCAGGCACTAAC
 DB 128 TTCACTGTACTCTGTGGCAAGGATGGTCCAGAAAGACCCCACTTCAGGCACTAAC
 QY 1213 GCTGGACCTGGCGCAGGAAGCAAGAGACTGGGCTTAGCCAGGAGTTCCTCCAA
 DB 188 GCTGGACCTGGCGCAGGAAGCAAGAGACTGGGCTTAGCCAGGAGTTCCTCCAA
 QY 1273 AGGGCGGAGAAACAGACAAAGCTCCTCCCTTGAGAAATTCCTGTGGATTTTAA
 DB 248 AGGGCGGAGAAACAGACAAAGCTCCTCCCTTGAGAAATTCCTGTGGATTTTAA
 QY 1333 TATTATTTTATTATTATTGTGACAAATGTTGATAATGG 1373
 DB 308 TATTATTTTATTATTATTGTGACAAATGTTGATAATGG 348

RESULT 27

BE858778 710 bp mRNA linear EST 2;
 BE858778/c LOCUS
 7f9db06.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3;
 DEFINITION
 similar to contains element MER32 repetitive element ; mi
 sequence.
 ACCESSION
 BE858778
 VERSION
 BE858778.1 GI:10374165
 KEYWORDS
 EST
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 710)
 REFERENCE
 NCI/NCI-NDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute / National Institute of Neurolo
 Disorders and Stroke, Brain Tumor Genome Anatomy Project
 (CGAP/BTGAP), Tumor Gene Index
 UNPUBLISHED (1998)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenf

Library Preparation: M. Bento Soares, Ph.D., M. Fatima Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
Sequencing by: Washington University Genome Sequencing Center
Distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
age.llnl.gov
mer: -40UP from Gibco
Library sequence stop: 342.
Location/Qualifiers
1. 710

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3304691"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/note="Organ: Brain; Vector: pTVT3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTACCAATCGAAGTGGAGCGGCATATCTTTTTTTTTTTTTTTTTTT
T 3'); double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pTVT3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

24.0%; Score 329; DB 10; Length 710;
rity 100.0%; Pred. No. 4e-156;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
CCCCACGCCCACTCTCCACTACTAGCTCCCAATCCCTGACCTTTGAGGCC 1104
CCCCACGCCCACTCTCCACTACTAGCTCCCAATCCCTGACCTTTGAGGCC 289
TGATCTGACTCCCCCTGGCCACAGACCCCGAGGCGATTGTCTACTGTACTC 1164
TGATCTGACTCCCCCTGGCCACAGACCCCGAGGCGATTGTCTACTGTACTC 229
ACAGGATGGTCCAGAGACCCCACTTACGCGCTTAAGAGGGGCTGACCTGGC 1224
ACAGGATGGTCCAGAGACCCCACTTACGCGCTTAAGAGGGGCTGACCTGGC 169
GAAGCCAAAGAGACTGGCGCTAGGCGAGGAGTCCCAATGTGAGGGCGAGAAA 1284
GAAGCCAAAGAGACTGGCGCTAGGCGAGGAGTCCCAATGTGAGGGCGAGAAA 109
CAAGCTCTCCCTTTCAGAAATCCCTGTGGATTCTTTAAACAGATATTTTAT 1344
CAAGCTCTCCCTTTCAGAAATCCCTGTGGATTCTTTAAACAGATATTTTAT 49
TTGTGACAAATTTGTGATAATGG 1373
TTGTGACAAATTTGTGATAATGG 20

367 bp mRNA linear EST 27-FEB-2002
11-acs-a-05-0-UI.81 UI-E-Q1 Homo sapiens cDNA clone
11-acs-a-05-0-UI 3', mRNA sequence.
2.1 GI:18967291
apiens (human)
apiens
ata; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 367)
M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PubMed 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of
Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Res
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-94, >At rich#Low complexity (matched complement
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1. 367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-Q1-acs-a-05-0-UI"
/tissue_type="optic nerve"
/dev_stage="adult"
/clone_lib="UI-E-Q1"
/note="Organ: eye; Vector: pTVT3-Pac (Pharmacia) v
modified polylinker; Site 1: EcoRI; Site 2: Not I
UI-E-Q1 is a normalized cDNA library containing t
following tissue(s): optic nerve. The library was
constructed according to Bonaldo, Lennon and Soar
Genome Research, 6:791-806, 1996. First strand cD
Synthesis was primed with an oligo-dT primer cont
Not I site. Double stranded cDNA was ligated to a
adaptor, digested with Not I, and cloned directio
into pTVT3-Pac vector. The oligonucleotide used t
the synthesis of first-strand cDNA contains a lib
sequence that is located between the Not I site a
(dT)18 tail. The sequence tag for this library is
CCATTAAGTG. This library was created for the prog
discovery in the Visual System, supported by Natio
Institute (NEI).
TAG_TISSUE=human optic nerve
TAG_LIB=UI-E-Q1
TAG_SEQ=CCATTAAGTG"

ORIGIN

Query Match 23.9%; Score 328; DB 12; Length 367;
Best Local Similarity 100.0%; Pred. No. 1.1e-155;
Matches 328; Conservative 0; Mismatches 0; Indels 0; G
QY 1045 ACTCCCCACCGCCACTCTCCACTACTAGCTCCCAATCCCTGACCTTTGAG
DB 367 ACTCCCCACCGCCACTCTCCACTACTAGCTCCCAATCCCTGACCTTTGAG
QY 1106 CCAGTGATCTGACTCCCTCCCTGGCCACAGACCCCGAGGCGATTGTCTACTGTAC
DB 307 CCAGTGATCTGACTCCCTCCCTGGCCACAGACCCCGAGGCGATTGTCTACTGTAC
QY 1166 GTGGGCAAGAGTGGGTCCAGAGAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTT
DB 247 GTGGGCAAGAGTGGGTCCAGAGAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTT
QY 1226 GCAGGAGCAAGAGACTGGGCGCTAGGCGAGGAGTCCCAATGTGAGGGGCGAG
DB 187 GCAGGAGCAAGAGACTGGGCGCTAGGCGAGGAGTCCCAATGTGAGGGGCGAG

CAAGCTCCTCCCTTGAGAAATCCCTGCGATTTTAAACAGATATATTTTATT 1345
|||||
CAAGCTCCTCCCTTGAGAAATCCCTGCGATTTTAAACAGATATATTTTATT 68
|||||
TTGTGACAAATGTTGATAAATGG 1373
|||||
TTGTGACAAATGTTGATAAATGG 40
|||||
22 346 bp mRNA linear EST 29-SEP-2000
4.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3308766 3',
sequence.
22
22.1 GI:10374253
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 346)
NDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
al Cancer Institute / National Institute of Neurological
ers and Stroke, Brain Tumor Genome Anatomy Project
BTGAP), Tumor Gene Index
ished (1998)
t: Robert Strausberg, Ph.D.
cgapbs-r@mail.nih.gov
Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Library Preparation: M. Bento Soares, Ph.D., M. Fatima
o, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
sequencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
mage.llnl.gov
amer: -40UP from Gibco
quality sequence stop: 344.
Location/Qualifiers
1. 346
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3308766"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn23"
/note="Organ: brain; Vector: pTT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pTT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
23.8%; Score 327; DB 10; Length 346;
arity 100.0%; Pred.No. 3.6e-155; Mismatches 0; Indels 0; Gaps 0;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
|CCCCACGGCCACTCTCCACTAGTCTCCCAATCCCTGACCCCTTTGAGGCCCC 1106
|||||
|CCCCACGGCCACTCTCCACTAGTCTCCCAATCCCTGACCCCTTTGAGGCCCC 287
|||||
|GATCTCGACTCCCTCGCCACAGACCCCGAGGCAATGTGTTCACTGTACTCTG 1166
|||||
|GATCTCGACTCCCTCGCCACAGACCCCGAGGCAATGTGTTCACTGTACTCTG 227
|||||
|CAAGATGGTCCAGNAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCG 1226
|||||

Db 226 TGGCAAGGATGGTCCGAAGACCCCACTTCAGGCACTAGAGGGGCTGGACCTG
|||||
Qy 1227 CAGGAAGCCAAAGAGACTGGGCGCTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGA
|||||
Db 166 CAGGAAGCCAAAGAGACTGGGCGCTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGA
|||||
Qy 1287 AGCAAGCTCCTCCCTTGAGAAATCCCTGCGATTTTAAACAGATATATTTT
|||||
Db 106 AGCAAGCTCCTCCCTTGAGAAATCCCTGCGATTTTAAACAGATATATTTT
|||||
Qy 1347 TTATTGTGACAAATGTTGATAAATGG 1373
|||||
Db 46 TTATTGTGACAAATGTTGATAAATGG 20
|||||
RESULT 30
AW195034/c
LOCUS
DEFINITION
x45912.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:26
mRNA sequence.
ACCESSION
AW195034
VERSION
AW195034.1 GI:6474026
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 337)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (T
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Mic
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencin
Clone distribution: NCI-CGAP clone distribution informati
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 330.
Location/Qualifiers
1. 337
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2696710"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid11"
/note="Organ: kidney; Vector: pTT73D-Pac (Pharma
a modified polylinker; Site 1: Not I; Site 2: Eco
Plasmid DNA from the normalized library NCI_CGAP
prepared, and ss circles were made in vitro. Fol
purification, this DNA was used as tracer in a s
hybridization reaction. The driver was PCR-ampli
from a pool of 5,000 clones made from the same l
(clones: 132376-132391, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares an
Fatima Bonaldo."
ORIGIN
Query Match 23.7%; Score 326; DB 10; Length 337;
Best Local Similarity 100.0%; Pred.No. 1.2e-154;
Matches 326; Conservative 0; Mismatches 0; Indels 0;
Qy 1048 TCCCCACGGCCACTCTCCACTAGTCTCCCAATCCCTGACCCCTTTGAGG
|||||
Db 337 TCCCCACGGCCACTCTCCACTAGTCTCCCAATCCCTGACCCCTTTGAGG
|||||

CTCGACTCCCTCCGACAGACCCAGGCGCATTTGTCTTCACTGTACTGT 1167
 |||||
 CTGACTCCCTCCGACAGACCCAGGCGCATTTGTCTTCACTGTACTGT 218
 |||||

AGATGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGC 1227
 |||||
 AGATGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGC 158
 |||||

3CCAAAGAGACTGGGCTAGCCAGGAGTCCCAAAATGTGAGGGGCGAGAAACAA 1287
 |||||

3CCAAAGAGACTGGGCTAGCCAGGAGTCCCAAAATGTGAGGGGCGAGAAACAA 98
 |||||

3CTCTCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTATTTTATTAT 1347
 |||||

3CTCTCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTATTTTATTAT 38
 |||||

TGACAAATGTTGATAATGG 1373
 |||||

TGACAAATGTTGATAATGG 12
 |||||

3 345 bp mRNA linear EST 29-NOV-2000
 4.X1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3406302 3'
 to contains element MS1 repetitive element 1, mRNA

3.1 GI:11452510

piens (human)

piens

ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

a; Eutheria; Primates; Catarrhini; Hominidae; Homo.

es 1 to 345)

DS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

1 Cancer Institute / National Institute of Neurological

rs and Stroke, Brain Tumor Genome Anatomy Project

TGAP), Tumor Gene Index

shed (1998)

: Robert Strausberg, Ph.D.

cgabbs-remail.nih.gov

Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

library Preparation: M. Bento Soares, Ph.D., M. Fatima

Ph.D.

ibrary Arrayed by: Greg Lennon, Ph.D.

quencing by: Washington University Genome Sequencing Center

distribution: NCI-CGAP clone distribution information can be

through the I.M.A.G.E. Consortium/LLNL, send email to:

age.llnl.gov

mer: -40UP from Gibco

ality sequence stop: 333.

Location/Qualifiers

1. 345

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3406302"

/tissue_type="glioblastoma (pooled)"

/lab_host="DH10B"

/clone_lib="NCI CGAP Brn23"

/notes="Organ: Brain; Vector: pT7T3D-pac (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo (dT) primer [5'

TGTACCAATCTGAATGGAGCGGCGCATATCTTTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

Library is normalized, and was constructed by Bento

Soares and M.Fatima Bonaldo."

Query Match 23.7%; Score 326; DB 10; Length 345;
 Best Local Similarity 100.0%; Pred. No. 1.2e-154;
 Matches 326; Conservative 0; Mismatches 0; Indels 0; G+

QY 1048 TCCGCCACCGCCACTCTCCACCTCCTAGCTCCCATCCCTGACCTTTGAGGCC
 |||||
 DB 345 TCCGCCACCGCCACTCTCCACCTCCTAGCTCCCATCCCTGACCTTTGAGGCC
 |||||

QY 1108 AGTGATCTCGACTCCCTCCCTGCGCACAGACCCAGGCGCATTTGTCTCACTGTACTC
 |||||

DB 285 AGTGATCTCGACTCCCTCCCTGCGCACAGACCCAGGCGCATTTGTCTCACTGTACTC
 |||||

QY 1168 GGGCAAGATGGTCCAGAACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC
 |||||

DB 225 GGGCAAGATGGTCCAGAACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGC
 |||||

QY 1228 AGGAAGCCAAAGAGACTGGGCTAGGCGCAGAGTTCCCAAAATGTGAGGGGCGAGAAA
 |||||

DB 165 AGGAAGCCAAAGAGACTGGGCTAGGCGCAGAGTTCCCAAAATGTGAGGGGCGAGAAA
 |||||

QY 1288 GACAAGCTCTCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATTATTTTAT
 |||||

DB 105 GACAAGCTCTCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATTATTTTAT
 |||||

QY 1348 TATTGTGACAAATGTTGATAATGG 1373
 |||||

DB 45 TATTGTGACAAATGTTGATAATGG 20
 |||||

RESULT 32

AW204512/c

LOCUS

DEFINITION

UI-H-B11-aei-f-08-0-UI-s1 NCI CGAP_Sub3 Homo sapiens cDNA c

IMAGE:2719622 3', mRNA sequence.

ACCESSION

AW204512

VERSION

AW204512.1 GI:6503984

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 352)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (C

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

The sequence contained an oligo-dT track that was present

oligonucleotide that was used to prime the synthesis of fl

strand cDNA and therefore this may represent a bonafide po

tail. cDNA Library Preparation: M.B. Soares Lab Clone dist

NCI-CGAP clone distribution information can be found throug

I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image.html The following repeti

elements were found in this cDNA sequence: 49-81,

>AT rich#low complexity

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1. 352

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2719622"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP Sub3"

/note="Vector: pT7T3D-pac (Pharmacia) with a modifi

polylinker; Site 1: Not I; Site 2: Eco RI; The

NCI CGAP Sub3 library is a subtracted library deri

the NCI CGAP Sub1 library, which is a subtracted

derived from BI. BI constitutes a mixture of 21

normalized or subtracted NCI CGAP libraries:

NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr28, NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kid5, NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid11, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8, NCI CGAP CUL1, NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lu5, NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6, NCI CGAP Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
 NCI CGAP Kid3 pool 1 : LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clonides 1323376-1323911, 1456008-1456775, 1500552-1502855); NCI CGAP Kids pool 1 : LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonides 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI CGAP Lu5 pool 1 : LLAM 3575-3582, 3851-3854 (IMAGE Clonides 1414920-1417991, 1520904-1522439); NCI CGAP GC4 pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonides 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI CGAP Pr22 pool 1 : LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonides 985608-986759, 1101192-1101959, 1217928-1220615); NCI CGAP Co10 pool 1 : LLAM 2644-2653, 2871-2872 (IMAGE Clonides 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described [Bonaldi, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.
 TAG_TISSUE=breast
 TAG_LIB=NCI CGAP_Br2
 TAG_SEQ=AAACC"

23.7%; Score 326; DB 10; Length 352;
 arity 100.0%; Pred. No. 1.2e-154; Indels 0; Gaps 0;
 conservative 0; Mismatches 0;

CCACGGCCACTCTCCACCTCAGTCTCCCAATCCCTGACCTTTGAGGCCCCC 1107
 CCACGGCCACTCTCCACCTCAGTCTCCCAATCCCTGACCTTTGAGGCCCCC 293
 ATCTCGACTCCCTCCCTGCCACAGACCCCGAGGCGATTGTTCACGTACTCTGT 1167
 ATCTCGACTCCCTCCCTGCCACAGACCCCGAGGCGATTGTTCACGTACTCTGT 233
 AAGATGGCTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC 1227
 AAGATGGCTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGC 173
 AGCCAAAGAGACTGGGCTTAGCCAGGAGTTCCTCCCAATGTGAGGGCGAGAAACA 1287
 AGCCAAAGAGACTGGGCTTAGCCAGGAGTTCCTCCCAATGTGAGGGCGAGAAACA 113
 AGCTCCTCCCTGAGAAATCCCTGTGATTTTAAACAGATATATTTTATTAT 1347
 AGCTCCTCCCTGAGAAATCCCTGTGATTTTAAACAGATATATTTTATTAT 53
 GTGCAAAATGTTGTAATGG 1373
 GTGCAAAATGTTGTAATGG 27
 GTGCAAAATGTTGTAATGG 27

320 364 bp mRNA linear EST 16-JAN-2000
 312-agn-b-05-0-UI.s1 NCI CGAP Sub4 Homo sapiens CDNA clone
 :2724801 3', mRNA sequence.
 320
 320.1 GI:6698256

sapiens (human)
 sapiens
 ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 364)
 NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (Tumor Gene Index)
 TITLE Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 The sequence contained an oligo-dT track that was present oligonucleotide that was used to prime the synthesis of fi strand cDNA and therefore this may represent a bonafide po tail. cDNA library Preparation: M.B. Soares Lab Clone dist NCI CGAP clone distribution information can be found throu I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html The following repet elements were found in this cDNA sequence: 61-93,
 >AT-rich#Low complexity
 Seq primer: M13 Forward
 POLYA=Yes.

Location/Qualifiers
 source
 1. 364
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2724801"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Sub4"
 /notes="Vector: p773D-Pac (Pharmacia) with a modi polylinker; Site 1: Not I; Site 2: Eco RI; The NCI CGAP Sub4 library is a subtracted library der the NCI CGAP Sub2 library which is a subtracted derived from the NCI CGAP Sub1 library, which is subtracted library derived from BI. BI constit mixture of 21 normalized or subtracted NCI CGAP libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kid5, NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid11, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8, NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lu5, NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP Brn25. These 21 libraries were pooled ar single-stranded DNA preparation of the resulting was used as a tracer in a subtractive hybridizat a driver whose composition is detailed below:
 NCI CGAP Kid3 pool 1 : LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clonides 1323376-1323911, 1456008-1456775, 1500552-1502855); NCI CGAP Kids I LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clor 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI CGAP Lu5 pool 1 : LLAM 3575-3582, 3851-3854 (Clonides 1414920-1417991, 1520904-1522439); NCI CG pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IM Clonides 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI CGAP Pr22 pool 1 : LLAM 2457-2759, 3062-3068 (IMAGE Clonides 985608-986759, 1101192-1101959, 1217928-1220615); NCI CGAP Co10 F LLAM 2644-2653, 2871-2872 (IMAGE Clonides 1057416-1144584-1145351). Subtraction was performed as pre described [Bonaldi, Lennon & Soares (1996): Norm and Subtraction: Two Approaches To Facilitate Ge Discovery. Genome Research 6, 791-806.]
 TAG_TISSUE=kidney
 TAG_LIB=NCI CGAP_Kids
 TAG_SEQ=ATTTC"

ORIGIN
 Query Match 23.7%; Score 326; DB 10; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.2e-154;
 Matches 326; Conservative 0; Mismatches 0; Indels 0;
 QY 1048 TCCCCCAGCCCACTCTCCACCTCAGTCTCCCAATCCCTGACCTTTGAGG
 DB 364 TCCCCCAGCCCACTCTCCACCTCAGTCTCCCAATCCCTGACCTTTGAGG

TCCTGACTCCCTCCCTGCGCACAGACCCCGAGGCAATGTGTCACTGTACTCTGT 1167
 TCCTGACTCCCTCCCTGCGCACAGACCCCGAGGCAATGTGTCACTGTACTCTGT 245
 AGATGGTCCAGAAAGACCCCACTTCAGGCACTTAAGAGGGGCTGGACCTGGCGGC 1227
 AGATGGTCCAGAAAGACCCCACTTCAGGCACTTAAGAGGGGCTGGACCTGGCGGC 185
 GCCAAAGAGACTGGGCTAGCGCAGGAGTCCCAAAATGTGAGGGGCGAGAAACAA 1287
 GCCAAAGAGACTGGGCTAGCGCAGGAGTCCCAAAATGTGAGGGGCGAGAAACAA 125
 CTCCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATATATTTTATTAT 1347
 CTCCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATATATTTTATTAT 65
 TGACAAATGTTGATAAATGG 1373
 TGACAAATGTTGATAAATGG 39

1 .xl NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2297856 3',
 1 sequence.

1.1 GI:5633396

piens (human)

piens

ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

a; Eutheria; Primates; Catarrhini; Hominidae; Homo.

es 1 to 399)

p <http://www.ncbi.nlm.nih.gov/ncicgap>.

1 Cancer Institute, Cancer Genome Anatomy Project (CGAP),

ene Index

shed (1997)

: Robert Strausberg, Ph.D.

cgaps-r@mail.nih.gov

Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Buck, M.D., Ph.D.

library Preparation: M. Bento Soares, Ph.D.

quencing by: Greg Lennon, Ph.D.

distribution: NCI-CGAP clone distribution information can be

through the I.M.A.G.E. Consortium/LLNL at:

.llnl.gov/bbrp/image/image.html

mer: -40UP from Gibco.

Location/Qualifiers

1. .399

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_image="IMAGE:2297856"

/lab_host="DH10B"

/clone_lib="NCI CGAP Kid11"

/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not 1; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI_CGAP Kid3 was

prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(cloneIDs 1322376-1323911, 1456007-1456775, and

1500552-1502855). Subtraction by Bento Soares and M.

Fatima Bonaldo. "

23.6%; Score 324; DB 9; Length 399;

ity 100.0%; Pred. No. 1.3e-153; Indels 0; Gaps 0;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1050 CCCACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTTGAGCCCTTGAGGCCCC
 DB 347 CCCACCGCCCACTCTCCACCTCACTAGCTCCCAATCCCTTGAGCCCTTGAGGCCCC
 QY 1110 TGATCTCGACTCCCTCCCTGGCCACAGACCCCGAGGCAATGTGTCACTGTACTCTG
 DB 287 TGATCTCGACTCCCTCCCTGGCCACAGACCCCGAGGCAATGTGTCACTGTACTCTG
 QY 1170 GCAAGGATGGTCCAGAAAGACCCCACTTCAGGCACTTAAGAGGGGCTGGACCTGGCGG
 DB 227 GCAAGGATGGTCCAGAAAGACCCCACTTCAGGCACTTAAGAGGGGCTGGACCTGGCGG
 QY 1230 GAAGCCAAAGAGACTGGGCTAGCGCAGGAGTCCCAAAATGTGAGGGGCGAGAAACA
 DB 167 GAAGCCAAAGAGACTGGGCTAGCGCAGGAGTCCCAAAATGTGAGGGGCGAGAAACA
 QY 1290 CAAGCTCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATATTTTATTATTA
 DB 107 CAAGCTCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATATTTTATTATTA
 QY 1350 TTGTGACAAATGTTGATAAATGG 1373
 DB 47 TTGTGACAAATGTTGATAAATGG 24

RESULT 35

BI966255

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .456

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_image="IMAGE:5672623"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS

BI966255 456 bp mRNA linear EST 12-
 ie72904.yl Melton Normalized Human Islet 4 N4-HIS 1 Homo sa
 cDNA clone IMAGE:5672623 5' similar to TR:043508 043508 TNF
 WEAK INDUCER OF APOPTOSIS ; mRNA sequence.

ACCESION

BI966255

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Other ESTs: ie72904.xl

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cam

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencir

Washington University Genome Sequencing Center For informat

obtaining a clone please contact: Juliana Brown

(brown@fas.harvard.edu) This sequence now available from th

consortium, for clone orders contact: info@image.llnl.gov

High quality sequence stop: 429.

Location/Qualifiers

1. .456

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_image="IMAGE:5672623"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS

/note="Organ: Pancreas; Vector: pSPOR1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

23.6%; Score 324; DB 12; Length 456;
arity 99.7%; Pred. No. 1.3e-153;
onservative 0; Mismatches 1; Indels 0; Gaps 0;
CCCTCGGGCCCGAGCTCGGCTCTGCGAGTGTCTGGGCTGTGGCCCTGGGGCCA 765
CCCTCGGGCCCGAGCTCGGCTCTGCGAGTGTCTGGGCTGTGGCCCTGGGGCCA 141
CCTCCCTCGGGATCCGACACCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCTC 825
CCTCCCTCGGGATCCGACACCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCTC 201
ACTTCGAGACTTTCAGGTTCACTGAGGGGCCCTGTCTCCACAGCTGCTCCAG 885
ACTTCGAGACTTTCAGGTTCACTGAGGGGCCCTGTCTCCACAGCTGCTCCAG 261
CCGGTCCCTTCGACAGCTCTCTGGGACACCCGGTCCCTCTGCCCCACCTTCAGCC 945
CCGGTCCCTTCGACAGCTCTCTGGGACACCCGGTCCCTCTGCCCCACCTTCAGCC 321
TTTGCTCAGACCTGCCCTCCCTCTAGAGGCTGCGCTGGGCTGTTCACGTGTTT 1005
TTTGCTCAGACCTGCCCTCCCTCTAGAGGCTGCGCTGGGCTGTTCACGTGTTT 381
CCCATATAATACAGTATCCCACTTTATCTTACAACTCCCGCCAGCCCACTCT 1065
CCCATATAATACAGTATCCCACTTTATCTTACAACTCCCGCCAGCCCACTCT 441
CTCACTAGCTC 1080
CTCACTAGCTC 456

65 367 bp DNA linear GSS 27-AUG-1998
4_A1_H03_MF CIT Approved Human Genomic Sperm Library D Homo
is genomic clone Plate=3054 Col=5 Row=0, genomic survey
ice.
.65
.65.1 GI:3471394
apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 367)
as, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
ice-tagged connectors: A sequence approach to mapping and
ing the human genome
Nat'l. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
589
764
ct: Mahairas GG, Wallace JC, Hood L
throughput Sequencing Center

University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3054 row: 0 column: 5
Class: BAC ends
High quality sequence stop: 367.
FEATURES
Location/Qualifiers
1..367
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3054 Col=5 Row=0"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Libr
/note="Organ: sperm; Vector: pBelobAC11; BAC Clon
E-Coli DH10B"

ORIGIN

Query Match 23.4%; Score 321; DB 28; Length 367;
Best Local Similarity 100.0%; Pred. No. 4.2e-152;
Matches 321; Conservative 0; Mismatches 0; Indels 0; G
QY 776 TGGGATCGGACACCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCTCCTACCTAC
DB 1 TGGGATCGGACACCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCTCCTACCTAC
QY 836 GACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCACAGCTGCTCCAGGCTGCC
DB 61 GACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTCCACAGCTGCTCCAGGCTGCC
QY 896 CCCCTCGACAGCTCTCTGGGACACCCGGTCCCTCTGCCCCACCTTCAGCCCTCTT
DB 121 CCCCTCGACAGCTCTCTGGGACACCCGGTCCCTCTGCCCCACCTTCAGCCCTCTT
QY 956 CCAGACCTGCCCCCTCCCTCTAGAGGCTGCGCTGGGCTGTTCACGTGTTTCCATCC
DB 181 CCAGACCTGCCCCCTCCCTCTAGAGGCTGCGCTGGGCTGTTCACGTGTTTCCATCC
QY 1016 TAATACAGTATCCCACTTTATCTTACAACTCCCGCCAGCCCACTCTCCACCT
DB 241 TAATACAGTATCCCACTTTATCTTACAACTCCCGCCAGCCCACTCTCCACCT
QY 1076 AGCTCCCCCAATCCCTGACCCCT 1096
DB 301 AGCTCCCCCAATCCCTGACCCCT 321

RESULT 37

BM925491
LOCUS 1819 bp mRNA linear EST 1;
DEFINITION AGENCOURT 6625205 NIH_MGC_114 Homo sapiens cDNA clone IMA
5', mRNA sequence.
ACCESSION BM925491
VERSION BM925491.1 GI:19375870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1819)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information at

rough the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov

LLAM12814 row: n column: 16

ality sequence start: 74

ality sequence stop: 420.

Location/Qualifiers

1. .1819

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5763279"

/lab_host="DH10B"

/clone_lib="NIH MGC 114"

/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

23.4%; Score 321; DB 12; Length 1819;

ality 100.0%; Pred. No. 5.9e-152;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

ICCACTCTCCACTCAGTCTCCCAATCCCTGACCCCTTTGAGGCCCCAGTGA 1112

ICCACTCTCCACTCAGTCTCCCAATCCCTGACCCCTTTGAGGCCCCAGTGA 139

ACTCCCCCTGGCCACAGCCCCCAGGGGCTGTCTCACTGTACTCTGTGGCA 1172

ACTCCCCCTGGCCACAGCCCCCAGGGGCTGTCTCACTGTACTCTGTGGCA 199

GGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGAGAA 1232

GGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGAGAA 259

AGAGACTGGGCTAGGCAGAGATCCCAATGTGAGGGGCGAGAACAGACAA 1292

AGAGACTGGGCTAGGCAGAGATCCCAATGTGAGGGGCGAGAACAGACAA 319

TCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTATTTTATTATTG 1352

TCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTATTTTATTATTG 379

AAATGTTGATAAATGG 1373

AAATGTTGATAAATGG 400

19 374 bp mRNA linear EST 14-FEB-2002

5.xl Human insulinoma Homo sapiens cDNA 3', mRNA sequence.

19 1 GI:18669065

apiens (human)

apiens

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ses 1 to 374)

D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

ka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

r, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

t, A., Theising, B., Ritter, B., Ronko, I., Bennett, J.,

as, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,

ms, T., Jackson, Y. and Bowers, Y.

ine Pancreas Consortium

ished (2000)

COMMENT

Other ESTs: ihl5b05.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Can
MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excis
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washing
University Genome Sequencing Center For information on obt
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.ec
Seq primer: -40UP from Gibco.

FEATURES

source

1. .374

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="insulinoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Human insulinoma"

/note="Organ: pancreas; Vector: pBluescript SK-;
XhoI; Site 2: EcoRI; Constructed with lambda ZAPI
pBluescript SK- by Dr. J. Ferrer, in vivo mass-excis
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml.
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division
Laboratory, Washington University School of Medic
8127, 660 S Euclid Ave, St. Louis, MO 63110). No
is a Washington University Pancreas EST project 1;

ORIGIN

Query Match 23.3%; Score 320; DB 12; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.4e-151; Indels 0; G
Matches 320; Conservative 0; Mismatches 0;

QY 1054 ACCGCCCACTCTCCACTCAGTCTCCCAATCCCTGACCCCTTTGAGGCCCCAG

Db 346 ACCGCCCACTCTCCACTCAGTCTCCCAATCCCTGACCCCTTTGAGGCCCCAG

QY 1114 CTGACTCCCCCTGGCCACAGACCCCGGCGCATTTGTTCTACTCTGTGG

Db 286 CTGACTCCCCCTGGCCACAGACCCCGGCGCATTTGTTCTACTCTGTGG

QY 1174 GATGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGGCTCGACTCGCGCAG

Db 226 GATGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGGCTCGACTCGCGCAG

QY 1234 CCAAGAGACTGGGCTTAGGCCAGGAGTCCCAATGTGAGGGGCGAGAACAGA

Db 166 CCAAGAGACTGGGCTTAGGCCAGGAGTCCCAATGTGAGGGGCGAGAACAGA

QY 1294 CTCTCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTATTTATTATTA

Db 106 CTCTCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTATTTATTATTA

QY 1354 GACAAATGTTGATAAATGG 1373

Db 46 GACAAATGTTGATAAATGG 27

RESULT 39

AI695776/c

LOCUS

DEFINITION

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

AI695776 329 bp mRNA linear EST 17

wb77907.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:231

AI695776

AI695776.1 GI:4983676

EST.

Homo sapiens (human)

apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 329)
AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
cgapbs-remail.nih.gov
Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
-Buck, M.D., Ph.D.
Library Preparation: M. Bento Soares, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
sequencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
o.llnl.gov/bbrp/image/image.html
Length: 305 Std Error: 0.00
imer: -40UP from Gibco
uality sequence stop: 302.
Location/Qualifiers
1. 329
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2311740"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: p7T3D-Pac (Pharmacia)
with a modified polylinker; plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and as
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones from the same library (cloneids
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

23.2%; Score 318; DB 9; Length 329;
arity 100.0%; Pred. No. 1.4e-150;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
CACTCTCCACCTCACTAGTCCCAATCCCTGACCCCTTGAGGCCCCAGTGATCT 1115
CACTCTCCACCTCACTAGTCCCAATCCCTGACCCCTTGAGGCCCCAGTGATCT 270
TCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGG 1175
TCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGG 210
TCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGG 1235
TCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGG 150
TCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGG 1295
TCCCCCTGGCCACAGACCCCCAGGGCATTGTGTTCACTGTACTCTGTGGCAAGG 90
CCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTTATTTATTTATTTATTTGTA 1355
CCCTTGAGAAATCCCTGTGGATTTTAAACAGATATTTATTTATTTATTTATTTGTA 30
AATGTTGATAAATGG 1373
AATGTTGATAAATGG 12

LOCUS BF195436 340 bp mRNA linear EST 03
DEFINITION 7n17g12.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:35
mRNA sequence.
ACCESSION BF195436
VERSION BF195436.1 GI:11082306
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 340)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurolog
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfe
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fati
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencin
Clone distribution: NCI-CGAP clone distribution informati
found through the I.M.A.G.E. Consortium/LLNL, send email t
info@image.llnl.gov.

FEATURES
Location/Qualifiers
1..340
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3564887"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/note="Organ: brain; Vector: p7T3D-Pac (Pharmac
modified polylinker; Site 1: Not I; Site 2: Eco I
strand cDNA was primed with a Not I - oligo(dT) I
TGTACCAATCTGAAGTGGAGCGCGGCATCTTTTTTTTTTTT
T 3'; double-stranded cDNA was ligated to Eco I
adaptors (Pharmacia), digested with Not I and c
the Not I and Eco RI sites of the modified p7T
Library is normalized, and was constructed by B
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 23.2%; Score 318; DB 10; Length 340;
Best Local Similarity 100.0%; Pred. No. 1.4e-150;
Matches 318; Conservative 0; Mismatches 0; Indels 0;
QY 1056 CGCCCACTCTCCACCTCACTAGTCCCAATCCCTGACCCCTTGAGGCCCCAGT
DB 340 CGCCCACTCTCCACCTCACTAGTCCCAATCCCTGACCCCTTGAGGCCCCAGT
QY 1116 CCACTCCCCCTGGCCACAGACCCCGAGGGCATTGTGTTCACTGTACTCTGTGGG
DB 280 CCACTCCCCCTGGCCACAGACCCCGAGGGCATTGTGTTCACTGTACTCTGTGGG
QY 1176 ATGGGTCCAGAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGAGG
DB 220 ATGGGTCCAGAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGAGG
QY 1236 AAAGAGACTGGGCCCTAGGCCAGAGTTCCTCAATGTGAGGGGCGAGAAACAAGAC
DB 160 AAAGAGACTGGGCCCTAGGCCAGAGTTCCTCAATGTGAGGGGCGAGAAACAAGAC
QY 1296 CCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATTTATTTATTTATTTAT
DB 100 CCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATTTATTTATTTATTTAT
QY 1356 CAAATGTTGATAAATGG 1373

IGTTGATAAATGG 23

5 .x1 Soares_NFL_T_GBC_S1 mRNA linear EST 30-NOV-1998
842906 3', mRNA sequence.

5.1 GI:3804188

piens (human)

piens

a; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ta; Eutheria; Primates; Catarrhini; Hominidae; Homo.

es 1 to 407

P <http://www.ncbi.nlm.nih.gov/ncicgap>.

1 Cancer Institute, Cancer Genome Anatomy Project (CGAP),

ene Index

shed (1997)

: Robert Strausberg, Ph.D.

cgapbs-r@mail.nih.gov

one is available royalty-free through LLNL; contact the
onsortium (info@image.llnl.gov) for further information.

Length: 610 Std Error: 0.00

mer: -40UP from Gibco

ality sequence stop: 401.

Location/Qualifiers

1. .407

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1842906"

/lab_host="DH10B"

/clone_lib="Soares_NFL_T_GBC_S1"

/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell

NCI CGAP GCBI) were mixed and ss circles were made in
vitro. Following HAP purification, this DNA was used as

tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made

from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by Bento

Soares and M. Fatima Bonaldo.

23.0%; Score 316; DB 9; Length 407;

urity 100.0%; Pred. No. 1.5e-149;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

TCTCCACCTCCTAGTCTCCCAATCCCTGACACCTTTGAGGCCCCCGAGTATCTCG 1117

TCTCCACCTCCTAGTCTCCCAATCCCTGACACCTTTGAGGCCCCCGAGTATCTCG 276

CCCTTGGCCACAGACCCCGAGGCGATGTGTTCACTGTACTCTGTGGGCAAGGAT 1177

CCCTTGGCCACAGACCCCGAGGCGATGTGTTCACTGTACTCTGTGGGCAAGGAT 216

CCAGAGACCCCACTTCAGGACACTTAAGAGGGGCTGGACCTTGGCGGCGAGGAGCCAA 1237

CCAGAGACCCCACTTCAGGACACTTAAGAGGGGCTGGACCTTGGCGGCGAGGAGCCAA 156

ACTGGGCTAGGCCAGAGTTCCCAATGTGAGGGGCGAGAAACAAGCAAGCTCC 1297

ACTGGGCTAGGCCAGAGTTCCCAATGTGAGGGGCGAGAAACAAGCAAGCTCC 96

TTGAGATTCCCTGTGATTTTAAAAACAGATATTATTTTATTATTATTGTGACA 1357

TTGAGATTCCCTGTGATTTTAAAAACAGATATTATTTTATTATTATTGTGACA 36

QY 1358 AAATGTTGATAAATGG 1373

Db 35 AAATGTTGATAAATGG 20

RESULT 42

A1291866/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

A1291866 416 bp mRNA linear EST 29-
qm86c02.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:18956
mRNA sequence.

A1291866 GI:3934640

EST.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Eutelec

1 (bases 1 to 416)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (C

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Mici

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution informati

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 649 Std Error: 0.00

Seg primer: -40UP from Gibco

High quality sequence stop: 411.

Location/Qualifiers

1. .416

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1895618"

/tissue_type="carcinoid"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Lu5"

/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia
modified polylinker; 1st strand cDNA was prepared

neuroendocrine lung carcinoid, and was then prime
Not I - oligo(dT) primer. Double-stranded cDNA wa

to Eco RI adaptors (Pharmacia), digested with Not
cloned into the Not I and Eco RI sites of the mod

pT7T3 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo

ORIGIN

Query Match 22.5%; Score 309; DB 9; Length 416;

Best Local Similarity 100.0%; Pred. No. 5.7e-146; Indels 0; G

Matches 309; Conservative 0; Mismatches 0; Indels 0; G

QY 1065 TCACCTCTACTAGTCTCCCAATCCCTGACCTTTGAGGCCCCCGAGTATCTCGACT

Db 334 TCACCTCTACTAGTCTCCCAATCCCTGACCTTTGAGGCCCCCGAGTATCTCGACT

QY 1125 CTGTGGCCACAGACCCCGAGGCGATTTGTGTTTCACTGTACTGTGGGCAAGGATGGG

Db 274 CTGTGGCCACAGACCCCGAGGCGATTTGTGTTTCACTGTACTGTGGGCAAGGATGGG

QY 1185 GAAGACCCCACTTCAGGCGCTAAGAGGGGCTGGACCTGGCGGCGAGGAAGCCAAAGA

Db 214 GAAGACCCCACTTCAGGCGCTAAGAGGGGCTGGACCTGGCGGCGAGGAAGCCAAAGA

QY 1245 GGGCCTAGGCCAGGAGTTCCCAATGTAGGGGCGAGAAACAAGCAAGCTCTCTCC

Db 154 GGGCCTAGGCCAGGAGTTCCCAATGTAGGGGCGAGAAACAAGCAAGCTCTCTCC

TTCCCTGTGATTTTAAACACATATTATTTTATTTATTTATTTGTCACAAATGTT 1364
 TTTCCCTGTGATTTTAAACACATATTATTTTATTTATTTATTTGTCACAAATGTT 35
 AATGG 1373
 AATGG 26
 53 667F1 NIH_MGC_86 910 bp mRNA linear EST 30-JAN-2001
 sequence. IMAGE:4367225 5',
 63 GI:12603569
 sapiens (human)
 ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ses 1 to 910
 C http://mgc.nci.nih.gov/
 al Institutes of Health, Mammalian Gene Collection (MGC)
 ished (1999)
 t: Robert Strausberg, Ph.D.
 cgapbs-remail.nih.gov
 Procurement: ATCC
 Library Preparation: Life Technologies, Inc.
 Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 equencing by: Incyte Genomics, Inc.
 distribution: MGC clone distribution information can be
 through the I.M.A.G.E. Consortium/LLNL at:
 /image.llnl.gov
 LLAM10019 row: e column: 18
 quality sequence stop: 493.
 Location/Qualifiers
 1..910
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4367225"
 /tissue_type="osteosarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_86"
 /note="Organ: Bone; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.533 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 22.4%; Score 307; DB 12; Length 910;
 arity 99.8%; Pred. No. 7.1e-145;
 conservative 0; Mismatches 0; Indels 1; Gaps 1;
 TACCTGAAGCTGGACTTGTGTGATGTGTGCGCCCTCGCTGCTGCGAGCA 686
 TACCTGAAGCTGGACTTGTGTGATGTGTGCGCCCTCGCTGCTGCGAGCA 60
 TCAGCCACTGCGGCGAGTTCCCTCGGGGCCACAGCTCCGCTTGCAGGTGTCTGG 746
 TCAGCCACTGCGGCGAGTTCCCTCGGGGCCACAGCTCCGCTTGCAGGTGTCTGG 120
 TTGGCCCTCGGCGAGGTCTCCCTCGCGATCCGACACCTCCCTCGGGCCCATCT 806
 TTGGCCCTCGGCGCA-GGTCTCCCTCGGATCCGACACCTCCCTCGGGCCCATCT 179
 3GCTGCCCCCTTCTCACTTCTCGAGTCTTCCAGGTTCACGTGGGGCCCTGTT 866
 3GCTGCCCCCTTCTCACTTCTCGAGTCTTCCAGGTTCACGTGGGGCCCTGTT 239

QY 867 CTCCCCACAGTGTCCAGGCTGCCGGTCCCTCGACAGCTCTCTGGGACCCGG
 Db 240 CTCCCCACAGTGTCCAGGCTGCCGGTCCCTCGACAGCTCTCTGGGACCCGG
 QY 927 CTCTGCCCCACCTCAGCGCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCT
 Db 300 CTCTGCCCCACCTCAGCGCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCT
 QY 987 GGGCGCTGTTCACTGTTTCCATCCACATAAATACAGTATCCCACTCTTATCTT
 Db 360 GGGCGCTGTTCACTGTTTCCATCCACATAAATACAGTATCCCACTCTTATCTT
 QY 1047 CTCCCCCA 1054
 Db 420 CTCCCCCA 427
 RESULT 44
 AI202121/c
 LOCUS
 DEFINITION gi52c03.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:18
 mRNA sequence.
 ACCSSION AI202121
 VERSION AI202121.1 GI:3754727
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 317)
 NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute / National Institute of Neurolog
 Disorders and Stroke, Brain Tumor Genome Anatomy Project
 (CGAP/BTGP), Tumor Gene Index
 Unpublished (1998)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfe
 Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatij
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencir
 Clone distribution: NCI-CGAP clone distribution informati
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1005 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 308.
 Location/Qualifiers
 1..317
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1860100"
 /tissue_type="anaplastic oligodendroglioma"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Brn25"
 /note="Organ: brain; Vector: pT7T3D-Pac (Pharmac
 modified polylinker; Site 1: Not I; Site 2: Eco I
 strand cDNA was primed with a Not I - oligo(dT) I
 TGTACCAATCTGAGTGGAGCGCGCCATAGTGTCTTTTCTTTT
 T 3'); double-stranded cDNA was ligated to Eco I
 adaptors (Pharmacia), digested with Not I and c.
 the Not I and Eco RI sites of the modified pT7T.
 Library is normalized, and was constructed by B.
 Soares and M. Fatima Bonaldo."
 22.1%; Score 303; DB 9; Length 317;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 6.3e-143;
 Matches 303; Conservative 0; Mismatches 0; Indels 0;
 ORIGIN

32F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766071 5',
sequence.

9 9.1 GI:13917716

piens (human)

Diens

ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.

es 1 to 587)

http://img.nci.nih.gov/.

1 Institutes of Health, Mammalian Gene Collection (MGC)

shed (1999)

: Robert Strausberg, Ph.D.

cgapbs@mail.nih.gov

Procurement: Louis M. Staudt, M.D., Ph.D.

library Preparation: Ling Hong/Rubin Laboratory

library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

quencing by: Incyte Genomics, Inc.

distribution: MGC clone distribution information can be

through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov

LLCM1625 row: p column: 08

ality sequence stop: 587.

Location/Qualifiers

1..587

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4766071"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_48"

/notes="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by Ling

Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

20.6%; Score 283; DB 12; Length 587;

urity 99.6%; Pred. No. 1.2e-132;

nservative 0; Mismatches 1; Indels 1; Gaps 1;

XCGCGCGGTCCCGCTCCCGGATCCCTCGGGTCCCGGATGGGGGGGGGTGAGG 84

XCGCGCGGTCCCGCTCCCGGATCCCTCGGGTCCCGGATGGGGGGGGGTGAGG 76

XACAGCCCGCCCGCCCGATGGCCCGCGTGGAGCCAGAGCGGAGGGGGCGCGG 144

XACAGCCCGCCCGCCCGATGGCCCGCGTGGAGCCAGAGCGGAGGGGGCGCGG 136

AGCGGGCACCAGCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 203

AGCGGGCACCAGCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 196

CTCGGCTCTGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 263

CTCGGCTCTGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 256

CCTGCCAGGAGGAGTGGTGGCAGAGGAGGACAGGACCCGTCGGAATGCC 323

CCTGCCAGGAGGAGTGGTGGCAGAGGAGGACAGGACCCGTCGGAATGCC 316

ACAGAGAGAGGAGGAGTGGTGGCAGAGGAGGACAGGACCCGTCGGAATGCC 383

ACAGAGAGAGGAGGAGTGGTGGCAGAGGAGGACAGGACCCGTCGGAATGCC 376

GCACCTTAAGCGCGGAAACACAGGGCTCGAAGAGCGGATCGAGCCCATTTATGAGT 443

Db 377 AAGTGCACCTTAAGCGCGGAAACACAGGGCTCGAAGAGCGATCGAGCCCATTTATGA

Qy 444 TCATCCAGCAGCTCGACAGGAGCGGAGCGCAGCAG 478

Db 437 TCATCCAGCAGCTCGACAGGAGCGGAGCGCAGCAG 471

RESULT 49

AF163779

LOCUS

DEFINITION

AF163779 Human Homo sapiens genomic clone BAC750E14, genom

sequence.

ACCESSION

AF163779

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1027)

AUTHORS

Cousin, P., Billotte, J., Chaubert, P. and Shaw, P.H.

TITLE

Physical map of 17p13 and the genes adjacent to p53

JOURNAL

Genomics 63 (1), 60-68 (2000)

MEDLINE

20130114

PUBMED

10662545

COMMENT

Contact: Shaw PH

Experimental Oncology

Institute of Pathology

Rue du Bugnon 25, Lausanne, VD 1011, Switzerland

sub_clone=AB2R Asc-BamHI PSL1160

Classes: BAC subclone.

Location/Qualifiers

1..1027

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/map="117p"

/clone="BAC750E14"

/clone_lib="Human"

ORIGIN

Query Match

19.7%; Score 271; DB 28; Length 1027;

Best Local Similarity 99.7%; Pred. No. 1.7e-126;

Mismatches 321; Conservative 0; Mismatches 1; Indels 0; G

Qy 1052 CCACCCCGCCACTCTCCACCTCAGTCTCCCAATCCCTGACCCCTTTAGGCCCC

Db 268 CCACCCCGCCACTCTCCACCTCAGTCTCCCAATCCCTGACCCCTTTAGGCCCC

Qy 1112 ATCTCGACTCCCGCTGGCCACAGACCCCGGCGCATTTGTTCACGTACTCTGT

Db 328 ATCTCGACTCCCGCTGGCCACAGACCCCGGCGCATTTGTTCACGTACTCTGT

Qy 1172 AAGGATGGGTTCAGAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGC

Db 388 AAGGATGGGTTCAGAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGC

Qy 1232 AGCCAAAGAGACTGGGCTAGGCCAGGAGTCCCAATGTGAGGGGCGAGAAACAA

Db 448 AGCCAAAGAGACTGGGCTAGGCCAGGAGTCCCAATGTGAGGGGCGAGAAACAA

Qy 1292 AGCTCTCCCTTCAGAAATTCCTCTGTGGAATTTTAAAAACAGATATTTTATTTAT

Db 508 AGCTCTCCCTTCAGAAATTCCTCTGTGGAATTTTAAAAACAGATATTTTATTTAT

Qy 1352 GTGACAAAATGTTGATAAATGG 1373

Db 568 GTGACAAAATGTTGATAAATGG 589

RESULT 50

A1760777/c

77 346 bp mRNA linear EST 24-JUN-1999
5.xl NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2398377 3',
sequence.
77
77.1 GI:5176444
apiens (human)
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 346)
AP http://www.ncbi.nlm.nih.gov/ncicgap.
al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
ished (1997)
t: Robert Strausberg, Ph.D.
cgapbs-remail.nih.gov
Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
-Buck, M.D., Ph.D.
Library Preparation: M. Bento Soares, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
sequencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
o.llnl.gov/bbrp/image/image.html
imer: -40UP from Gibco.
Location/Qualifiers
1..346
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2398377"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid12"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid5 was
prepared, and ss circles were made in vitro following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."
19.4%; Score 267; DB 9; Length 346;
latity 99.7%; Pred. No. 1.5e-124;
onservative 0; Mismatches 1; Indels 0; Gaps 0;
CACTTCCACCTCAGTCTCCCAATCCCTGACCCCTTGGAGCCGCCAGTGATCT 1115
CACTTCCACCTCAGTCTCCCAATCCCTGACCCCTTGGAGCCGCCAGTGATCT 287
CTCCCCCTGGCCACAGACCCAGGCGCATTTGTTCACTGTACTCTGTGGCAAGG 1175
CTCCCCCTGGCCACAGACCCAGGCGCATTTGTTCACTGTACTCTGTGGCAAGG 227
GGTCCAGAGACCCCACTTCAAGGCACTTAAGAGGGGCTGGACCTGGCGGCAAGGCC 167
GAGACTGGGCTAGCCAGGAGTTCCTCAATGTGAGGGGCGAGAAACAACAGAGCT 1295
GAGACTGGGCTAGCCAGGAGTTCCTCAATGTGAGGGGCGAGAAACAACAGAGCT 107
CCCTTGAGAAATTCCTGTGATTTTAAACACAGATATTTATTTATTTATTTGTA 1355
CCCTTGAGAAATTCCTGTGATTTTAAACACAGATATTTATTTATTTATTTGTA 47
AATCTGATAAATGG 1373

Db 46 CAAAATGTTGATAAATGG 29
RESULT 51
BF940141/c
LOCUS
DEFINITION
BF940141 346 bp mRNA linear EST 22
nac68g06.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3
similar to contains Element MSRI repetitive element ;, mRN
sequence.
BF940141
BF940141.1 GI:12357461
EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
1 (bases 1 to 346)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurolog
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/RTGAP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfe
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fati
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencin
Distribution: NCI-CGAP clone distribution informati
found through the I.M.A.G.E. Consortium/LLNL, send email t
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 321.
Location/Qualifiers
1..346
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3439667"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn23"
/note="Organ: brain; Vector: pT73D-Pac (Pharmac
modified polylinker; Site 1: Not 1; Site 2: Eco I
strand cDNA was primed with a Not I - oligo(dT) I
TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTT
T 3'; double-stranded cDNA was ligated to Eco I
adaptors (Pharmacia), digested with Not I and c
the Not I and Eco RI sites of the modified pT7T
Library is normalized, and was constructed by B
Soares and M.Fatima Bonaldo."
ORIGIN
Query Match 19.1%; Score 262; DB 10; Length 346;
Best Local Similarity 100.0%; Pred. No. 5.4e-122; Indels 0;
Matches 262; Conservative 0; Mismatches 0;
QY 1048 TCCCCCAGCCGACCTCTCCACCTCACTAGCTCCCCCAATCCCTGACCCCTTGGG
346 TCCCCCAGCCGACCTCTCTCCACCTCACTAGCTCCCCCAATCCCTGACCCCTTGGG
1108 AGTGATCTCGACTCCCTCCCTGGCCACAGACCCCGGCGCATTTGTTCACTGTAC
286 AGTGATCTCGACTCCCTCCCTGGCCACAGACCCCGGCGCATTTGTTCACTGTAC
1168 GGGCAAGGATGGGTTCAGAAAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTG
226 GGGCAAGGATGGGTTCAGAAAGACCCCACTTCAGGCACTAAGAGGGGCTGACCTG
1228 AGGAAGCCAAAGAGACTGGGCTAGCCAGGAGTTCCTCAATGTGAGGGGCGGAGA

3CCAAAGAGACTGGCGCTAGCCAGGAGTTCCTCCAAATGTGAGGGCGGAGAAACAA 107

3CTCCCTCCCTTGAGAAAT 1309

3CTCCCTCCCTTGAGAAAT 85

8 561 bp mRNA linear EST 25-JAN-2001
2.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3441742 3'
to TR:Q9UK76 Q9UK76 HN1 PROTEIN. ; mRNA sequence.

8.1 GI:12512043

piens (human)

piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.

es 1 to 561)

IDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

rs and Stroke, Brain Tumor Genome Anatomy Project

TCAP), Tumor Gene Index

shed (1998)

: Robert Strausberg, Ph.D.

cgapbs-r@mail.nih.gov

Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

library Preparation: M. Bento Soares, Ph.D., M. Fatima

, Ph.D.

library Arrayed by: Greg Lennon, Ph.D.

quencing by: Washington University Genome Sequencing Center

distribution: NCI-CGAP clone distribution information can be

through the I.M.A.G.E. Consortium/LLNL, send email to:

age.llnl.gov

ality sequence stop: 260.

Location/Qualifiers

1. 561

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3441742"

/tissue_type="glioblastoma (pooled)"

/lab_host="DH10B"

/clone_lib="NCI CGAP Brn23"

/notes="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT73 vector.

Library is normalized, and was constructed by Bento

Soares and M. Fatima Bonaldo."

18.0%; Score 247; DB 10; Length 561;

arity 100.0%; Pred.No. 2.7e-114;

onservative 0; Mismatches 0; Indels 0; Gaps 0;

CACAGACCCCGAGGCAATGTGTTCACTGACTCTGTGGCGAGGATGGGTCCAGA 1186

CACAGACCCCGAGGCAATGTGTTCACTGACTCTGTGGCGAGGATGGGTCCAGA 212

CCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGAGGAGCAAGAGACTGG 1246

CCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGAGGAGCAAGAGACTGG 152

AGGCCAGGAGTTCCTCCAAATGTGAGGGCGGAGAAACAAGCAAGCTCTCCCTTGG 1306

AGGCCAGGAGTTCCTCCAAATGTGAGGGCGGAGAAACAAGCAAGCTCTCCCTTGG 92

QY 1307 AATTCCTGTGGATTTTAAACAGATATTATTTTATTTATTTATTTGTGACAAATGTI

Db 91 AATTCCTGTGGATTTTAAACAGATATTATTTTATTTATTTATTTGTGACAAATGTI

QY 1367 TAAATGG 1373

Db 31 TAAATGG 25

RESULT 53

AW081731/c

LOCUS

DEFINITION

AW081731 318 bp mRNA linear EST 14-
xb70a02.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
IMAGE:2581610 3', similar to contains element MSRI repetitiv
element ; mRNA sequence.

ACCESSION

AW081731 GI:6036883

VERSION

AW081731.1 Homo sapiens (human)

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleoc

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 318)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project ((

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact

IMAGE Consortium (info@image.llnl.gov) for further inform

Seq primer: -40UP from Gibco

High quality sequence stop: 314.

Location/Qualifiers

1. 318

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2581610"

/lab_host="DH10B"

/clone_lib="Soares NFL T GBC S1"

/notes="Organ: pooled; Vector: pT73D-Pac (Pharmac

a modified polylinker; Site 1: Not I; Site 2: Eco

Equal amounts of plasmid DNA from three normalize

libraries (fetal lung MbHL19W, testis NHT, and B-

NCI CGAP GCBI) were mixed, and ss circles were ma

vitro. Following HAP purification, this DNA was u

tracer in a subtractive hybridization reaction. T

was PCR-amplified cDNAs from pools of 5,000 clone

from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,

726408-728711, and 729096-731399. Subtraction by

Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 17.2%; Score 236; DB 9; Length 318;

Best Local Similarity 100.0%; Pred.No. 9.9e-109;

Matches 236; Conservative 0; Mismatches 0; Indels 0; G

QY 1138 CCCAGGCAATGTGTTCACTGACTCTGTGGCGAGGATGGGTCCAGAGACCCC

Db 254 CCCAGGCAATGTGTTCACTGACTCTGTGGCGAGGATGGGTCCAGAGACCCC

QY 1198 CAGCCTAAGAGGGGCTGGACCTGGCGGAGGAGCAAGAGACTGGGCTAGG

Db 194 CAGCCTAAGAGGGGCTGGACCTGGCGGAGGAGCAAGAGACTGGGCTAGG

QY 1258 GAGTTCCTCCAAATGTGAGGGCGGAGAAACAAGCAAGCTCTCTTGGAGATTC

Db 134 GAGTTCCTCCAAATGTGAGGGCGGAGAAACAAGCAAGCTCTCTTGGAGATTC

QY 1318 GATTTTAAACAGATATTATTTTATTTATTTATTTGTGACAAATGTTGATAATGG

|||||TTTAAACAGATATTATTTTATTTATTTATTTCTGACAAAATGTTGATAAATGG 19
 41 264 bp mRNA linear EST 06-APR-2000
 0.xl Soares NFL T.GBC.S1 Homo sapiens cDNA clone
 2978587 3', mRNA sequence.
 41 GI:7454367
 sapiens (human)
 apiens
 Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 es 1 to 264
 AP http://www.ncbi.nlm.nih.gov/ncicgap.
 al Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Gene Index
 ished (1997)
 ct: Robert Strausberg, Ph.D.
 : cgaps-r@mail.nih.gov
 : One is available royalty-free through LNL ; contact the
 Consortium (info@image.llnl.gov) for further information.
 umer: -40UP from Gibco
 uality sequence stop: 263.
 Location/Qualifiers
 1. .264
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2978587"
 /lab_host="DH10B"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19W, testis NHT, and B-cell
 NCI CGAP_GCB1) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

16.8%; Score 230; DB 10; Length 264;
 larity 100.0%; Pred. No. 1.1e-105;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ATTGTGTTCTACTGTCTGTGGCAAGGATGGTCCAGAGACCCCACTTCAGGCA 1203
 ATTGTGTTCTACTGTCTGTGGCAAGGATGGTCCAGAGACCCCACTTCAGGCA 195
 AGAGGGCTGGACCTGGCGGAGAACCCAAAGAGACTGGGCTTAGCCAGGAGTTC 1263
 AGAGGGCTGGACCTGGCGGAGAACCCAAAGAGACTGGGCTTAGCCAGGAGTTC 135
 AATGTGAGGGGCGAGAACCAAGCAAGCTCTCCCTTTGAGAATTCCTGTGGATTTT 1323
 AATGTGAGGGGCGAGAACCAAGCAAGCTCTCCCTTTGAGAATTCCTGTGGATTTT 75
 AACAGATATTTATTTATTTATTTATTTGACAAAATGTTGATAAATGG 1373
 AACAGATATTTATTTATTTATTTATTTATTTGACAAAATGTTGATAAATGG 25

563 253 bp mRNA linear EST 21-DEC-1999

DEFINITION w17905.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE.24
 similar to contains element PTR5 repetitive element ;, mRNA
 sequence.
 ACCSSION AI863563
 VERSION AI863563.1 GI:5527670
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 253)
 NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Mic
 Emert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution informati
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 669 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 253.
 Location/Qualifiers
 1. 253
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2403128"
 /tissue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP_Kid12"
 /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharma
 a modified polylinker; Site 1: Not 1; Site 2: Eco
 plasmid DNA from the normalized library NCI CGAP
 prepared, and ss circles were made in vitro. Fol
 purification, this DNA was used as tracer in a s
 hybridization reaction. The driver was PCR-ampli
 from a pool of 5,000 clones made from the same 1
 (cloneIDs 1323912-1325831, 1471368-1472903 and
 1492104-1493255). Subtraction by Bento Soares an
 Fatima Bonaldo."

ORIGIN

Query Match 16.7%; Score 229; DB 9; Length 253;
 Best Local Similarity 100.0%; Pred. No. 3.5e-105;
 Matches 229; Conservative 0; Mismatches 0; Indels 0;
 QY 1145 GCATTGTGTTCTACTGTCTGTGGCAAGGATGGTCCAGAGACCCCACTTC
 Db 253 GCATTGTGTTCTACTGTCTGTGGCAAGGATGGTCCAGAGACCCCACTTC
 QY 1205 TAAGAGGGGCTGGACCTGGCGGAGAACCCAAAGAGACTGGGCTTAGCCAGG
 Db 193 TAAGAGGGGCTGGACCTGGCGGAGAACCCAAAGAGACTGGGCTTAGCCAGG
 QY 1265 CAAATGTGAGGGGCGAGAACCAAGCAAGCTCTCCCTTTGAGAATTCCTGTGGA
 Db 133 CAAATGTGAGGGGCGAGAACCAAGCAAGCTCTCCCTTTGAGAATTCCTGTGGA
 QY 1325 AAAACAGATATTTATTTATTTATTTATTTGACAAAATGTTGATAAATGG 1373
 Db 73 AAAACAGATATTTATTTATTTATTTATTTGACAAAATGTTGATAAATGG 25

RESULT 56

AI682487/c 238 bp mRNA linear EST 1
 LOCUS AI682487

apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ses 1 to 351)
NDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
al Cancer Institute / National Institute of Neurological
ers and Stroke, Brain Tumor Genome Anatomy Project
BTGAP), Tumor Gene Index
ished (1998)
t: Robert Strausberg, Ph.D.
cgaps-remail.nih.gov
Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Library Preparation: M. Bento Soares, Ph.D., M. Fatima
o, Ph.D.
Library Arrayed by: Greg Lennon, Ph.D.
quencing by: Washington University Genome Sequencing Center
distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL, send email to:
mage.llnl.gov
imer: -40UP from Gibco
uality sequence stop: 339.
Location/Qualifiers
1. .351
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3441693"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/note="Organ: brain; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCATCTGAGTGGAGCGCGCCGATATCTTTTATTTTATTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T3D vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

15.4%; Score 211; DB 10; Length 351;
arity 99.6%; Pred. No. 5.8e-96;
onservative 0; Mismatches 1; Indels 0; Gaps 0;
CCACGCCCACTCTCCACTCTAGCTCCCAATCCCTGACCTTTGAGGCCCCC 1107
CCACGCCCACTCTCCACTCTAGCTCCCAATCCCTGACCTTTGAGGCCCCC 292
ATCTCGACTCCCTCCCTGGCCACAGACCCCGAGGCATTGTGTCACTGTACTCTGT 1167
ATCTGACTCCCTCCCTGGCCACAGACCCCGAGGCATTGTGTCACTGTACTCTGT 232
AAGGATGGTCCAGAGACCCCACTTACGGCACTAAGAGGGCTGACCTGGCGC 1227
AAGGATGGTCCAGAGACCCCACTTACGGCACTAAGAGGGCTGACCTGGCGC 172
AAGCAAGAGACTGGCCCTAGCCAGGAGTCCCAATGTGAGGGCGGAGAACAA 1287
AAGCAAGAGACTGGCCCTAGCCAGGAGTCCCAATGTGAGGGCGGAGAACAA 112
AAGCTCCCTCCCTTGAGAAAT 1309
AAGCTCCCTCCCTTGAGAAAT 90

243 397 bp mRNA linear EST 17-DEC-1999
01.xl NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2315088 3',
sequence.
243

AI69243.1 GI:4834017
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 397)
REFERENCE
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (Tumor Gene Index)
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Mi
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencin
Clone distribution: NCI-CGAP clone distribution informati
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 438 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .397
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2315088"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr28"
/note="Organ: prostate; Vector: p7T3D-Pac (Pharm
with a modified polylinker; plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, an
circles were made in vitro. Following HAP purific
this DNA was used as tracer in a subtractive hybr
reaction. The driver was PCR-amplified cDNAs fro
of 5,000 clones made from the same library (clon
985608-986759, 110192-1101959, and 1217928-1220
Subtraction by Bento Soares and M. Fatima Bonald

Query Match 15.1%; Score 207; DB 9; Length 397;
Best Local Similarity 99.4%; Pred. No. 6.6e-94;
Matches 307; Conservative 0; Mismatches 2; Indels 0;
QY 1065 TCCACCTCACTAGCTCCCAATCCCTGACCTTTGAGGCCCCCAGTGTCTCGAC'
DB 343 TCCACCTCACTAGCTCCCAATCCCTGACCTTTGAGGCCCCCAGTGTCTCGAC'
QY 1125 CTTGGCCACAGACCCCGAGGCATTGTGTCACTGTACTCTGTGGCAAGGATGG
DB 283 CTTGGCCACAGACCCCGAGGCATTGTGTCACTGTACTCTGTGGCAAGGATGG
QY 1185 GAAGACCCCACTTACGGCACTAAGAGGGCTGGA CTTGGCGGCGAGGCCAAAG
DB 223 GAAGACCCCACTTACGGCACTAAGAGGGCTGGA CTTGGCGGCGAGGCCAAAG
QY 1245 GGGCCCTAGCCAGGAGTTC CCAATGTGAGGGCGGAGAACCAAGCAAGCTCCTC
DB 163 GGGCCCTAGCCAGGAGTTC CCAATGTGAGGGCGGAGAACCAAGCAAGCTCCTC
QY 1305 AGAATTCCTGTGGATTTTAAACAGATATTTATTTATTTATTTATTTGTGACAAA
DB 103 AGAATTCCTGTGGATTTTAAACAGATATTTATTTATTTATTTATTTGTGACAAA
QY 1365 GATAAATGG 1373
DB 43 GATAAATGG 35

```

4 26F1 NTH_MGC_118 Homo sapiens cDNA clone IMAGE:5217367 5',
sequence.
4 4.1 GI:16171193
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 894)
http://mgs.nci.nih.gov/
1 Institutes of Health, Mammalian Gene Collection (MGC)
shed (1999)
: Robert Strausberg, Ph.D.
cgabs@mail.nih.gov
Procurement: Life Technologies, Inc.
Library Preparation: Life Technologies, Inc.
Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Incyte Genomics, Inc.
Distribution: MGC clone distribution information can be
through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov
LLM11546 row: d column: 08
ality sequence start: 5
ality sequence stop: 460.
Location/Qualifiers
1. 894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5217367"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."
15.0%; Score 206; DB 12; Length 894;
urity 100.0%; Pred. No. 2.5e-93;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
AGGACGAGACCGCTCGGAACCTGAATCCCGACAGAAAGCCAGGATCTCTGG 351
AGGACGAGACCGCTCGGAACCTGAATCCCGACAGAAAGCCAGGATCTCTGG 281
CTCTGAACGACTAGTTCGCGCTCGCAGAGTGACCTTAAGCGCGGAAACACGG 411
CTCTGAACGACTAGTTCGCGCTCGCAGAGTGACCTTAAGCGCGGAAACACGG 341
AAGAGCGATCGCAGCGCCATTATGAAGTTTCATCCAGCAGCTGGACAGGACGAGCG 471
AAGAGCGATCGCAGCGCCATTATGAAGTTTCATCCAGCAGCTGGACAGGACGAGCG 401
AGGTGTGACGGGACAGTGAG 497
AGGTGTGACGGGACAGTGAG 427
41 5.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
EST 18-AUG-1998

```

```

IMAGE:1651448 3' similar to contains MSRI.t3 MSRI repetitive
element ;, mRNA sequence.
ACCESSION AI091441
VERSION AI091441.1 GI:3430500
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 465)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (C
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
This clone is available royalty-free through LLNL; contact
IMAGE Consortium (info@image.llnl.gov) for further informat
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 465.
Location/Qualifiers
1. 465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1651448"
/lab_host="DH10B"
/clone_lib="Soares NSF F8_9W_OT_PA_P_S1"
/notes="Organ: pooled; Vector: p773D-Pac (Pharmaci
a modified polylinker; Site 1: Not I; Site 2: Eco
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in
Following HAP purification, this DNA was used as t
a subtractive hybridization reaction. The driver v
PCR-amplified cDNAs from pools of 5,000 clones ma
the same 5 libraries. The pools consisted of the
libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 1500:
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares
pool 1: 723720-726407, 739080-740999 Subtraction
Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 14.7%; Score 202; DB 9; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.4e-91;
Matches 202; Conservative 0; Mismatches 0; Indels 0; G
QY 25 TCGGCGCGCGGCTCCCGCTCCCGGATCCCTCCCGGATCCCGGATCGGGGCGG
Db 183 TCGGCGCGCGGCTCCCGCTCCCGGATCCCTCCCGGATCGGGGCGG
QY 85 CAGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db 243 CAGGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
QY 145 GGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT
Db 303 GGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT
QY 205 TGCTCGGCGCTCTCTGCTGGCGG 226
Db 363 TGCTCGGCGCTCTCTGCTGGCGG 384
RESULT 62
BF222608/c
LOCUS BF222608
DEFINITION 224 bp mRNA linear EST 09
7p56d12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:364
mRNA sequence.
ACCESSION BF222608

```


Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Funding by: Agencourt Bioscience Corporation
 Distribution: MGC clone distribution information can be
 through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov
 NDAM593 row: e column: 12
 ality sequence stop: 249.
 Location/Qualifiers
 1..1064
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30520311"
 /tissue_type="Human Placenta"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC_147"
 /note="Organ: Placenta; Vector: pBluescriptR; Site: 1:
 all-XhoI; Site 2: BamH; Oligo-dr primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.3 kb and normalized to ROT 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci,
 preparation). Library constructed by M. Brownstein
 (NIMH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

12.7%; Score 175; DB 14; Length 1064;
 rity 100.0%; Pred.No. 1.7e-77;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 'CCCCCGCCCATGCGCCCGCTCGAGCCAGAGCGGGGGGGGA 149
 'CCCCCGCCCATGCGCCCGCTCGAGCCAGAGCGGGGGGGGA 87
 'GCACCGCCCTGTGTGCTCGCGCTCGCGCTGGGCTGGGCTGGGCTGGCT 209
 'GCACCGCCCTGTGTGCTCGCGCTCGCGCTGGGCTGGGCTGGGCTGGCT 147
 'TCCTGCTGGCGCTGTGTGCTGGGAGCGGGGATCGTGTCCGCCAG 264
 'TCCTGCTGGCGCTGTGTGCTGGGAGCGGGGATCGTGTCCGCCAG 202

13 1319 bp mRNA linear EST 26-SEP-2003
 RT 15623743 NIH_MGC_147 Homo sapiens cDNA clone
 10527869 5', mRNA sequence.
 13.1 GI:36348525
 piens (human)
 piens
 ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 es 1 to 1319)
 'http://mgi.nci.nih.gov/
 al Institutes of Health, Mammalian Gene Collection (MGC)
 shed (1999)
 t: Daniela S. Gerhard, Ph.D.
 of Cancer Genomics
 al Cancer Institute / NIH
 31 Rm10A07 Bethesda, MD 20892
 cgabs-remail.nih.gov
 Procurement: Dr. Stefan Hansson
 Library Preparation: Michael J. Brownstein (NHGRI) with help
 vice from Piero Carninci (RIKEN)
 Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 sequencing by: Agencourt Bioscience Corporation
 distribution: MGC clone distribution information can be
 through the I.M.A.G.E. Consortium/LLNL at:
 /image.llnl.gov

Plate: NDAM612 row: o column: 14
 High quality sequence stop: 287.
 Location/Qualifiers
 1..1319

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30527869"
 /tissue_type="Human Placenta"
 /lab_host="DH10B Tona"
 /clone_lib="NIH_MGC_147"
 /note="Organ: Placenta; Vector: pBluescriptR; Site
 all-XhoI; Site 2: BamH; Oligo-dr primed using prim
 5'-TTTTTTTTTTTTTTVN-3', size-selected for avera
 insert size 2.3 kb and normalized to ROT 5. This i
 primary library enriched for full-length clones ar
 constructed using the Cap-trapper method (Carninc
 preparation). Library constructed by M. Brownstei
 (NIMH/NHGRI, National Institutes of Health). Note:
 a NIH_MGC library."

ORIGIN

Query Match 12.7%; Score 175; DB 14; Length 1319;
 Best Local Similarity 100.0%; Pred.No. 1.8e-77;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; G
 QY 90 ACAGCCCCCGCCCATGCGCCCGCTCGAGCCAGAGCGGGGGGGGGG
 DB 29 ACAGCCCCCGCCCATGCGCCCGCTCGAGCCAGAGCGGGGGGGGGG
 QY 150 GCGGGGACCGCCCTGTGTGCTCGCGCTGGGCTGGGCTGGGCTGGGCT
 DB 89 GCGGGGACCGCCCTGTGTGCTCGCGCTGGGCTGGGCTGGGCTGGGCT
 QY 210 CGGCTCTCTGCTGGCGCTGTGTGCTGGGAGCGGGGATCGCTCTCGGCCAG
 DB 149 CGGCTCTCTGCTGGCGCTGTGTGCTGGGAGCGGGGATCGCTCTCGGCCAG

RESULT 66

AA913913 338 bp mRNA linear EST 24
 OL35h12.s2 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 IMAGE:1525511 3', mRNA sequence.

AA913913 GI:3053305

EST. Homo sapiens (human)

VERSION Homo sapiens

KEYWORDS Homo sapiens

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 338)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (

TITLE Unpublished (1997)

JOURNAL Tumor Gene Index

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

This clone is available royalty-free through LLNL; contac

IMAGE Consortium (info@image.llnl.gov) for further informa

Insert Length: 330 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 332.

Location/Qualifiers

1..338

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1525511"

/lab_host="DH10B"

/clone_lib="Soares_NFL_T_GBC_S1"

/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmac

a modified polylinker; Site_1: Not 1; Site_2: Eco

Equal amounts of plasmid DNA from three normalized libraries (fetal lung NDHL9W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made *in vitro*. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729036-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

12.7%; Score 174; DB 9; Length 330;
arity 99.7%; Pred. No. 4.4e-77;
onservative 0; Mismatches 0; Indels 1; Gaps 1;

CAATCCCTGACCCCTTGAGGGCCCCCAGTGTCTGCAGTCCGCCCTGGCCACAGACC 1138
CAATCCCTGACCCCTTGAGGGCCCCCAGTGTCTGCAGTCCGCCCTGGCCACAGACC 248
GGGCAATTGTGTCTCACTGTCTGTGGCAGAGTAGGGTCCAGAGAGACCCCACTTC 1198
GGGCAATTGTGTCTCACTGTCTGTGGCAGAGTAGGGTCCAGAGAGACCCCACTTC 188
ACTAAGAGGGGCTCGACTCTGGCGGAGGAAGCCTGGCCCTAGGCCAGG 1258
ACTAAGAGGGGCTCGACTCTGGCGGAGGAAGCCTAGGCCAGG 129
CCCAATTGTAGGGGCGAGAAACAGACAAGCTCCTCCTTGAGATTCCCTGTGG 1318
CCCAAAATGTAGGGGCGAGAAACAGACAAGCTCCTCCTTGAGATTCCCTGTGG 69
TTAAAAACAGATTATTTTTTATTATTATTGTGACAAAATGTTGATAAATGG 1373
TTAAAAACAGATTATTTTTTATTATTATTGTGACAAAATGTTGATAAATGG 14

97 196 bp mRNA linear EST 04-AUG-2000
12.x1 NCI CGAP_GC6 Homo sapiens cDNA clone IMAGE:3221115 3',
sequence.
197
197.1 GI:9703605
sapiens (human)
sapiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ases 1 to 196)
ZAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
nal Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Gene Index
lished (1997)
ct: Robert Strausberg, Ph.D.
c: cgaps-r@mail.nih.gov
a Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
ert-Buck, M.D., Ph.D.
Library preparation: M. Bento Soares, Ph.D., M. Fatima
do. Ph.D.

Library Arrayed by: Greg Lennon, Ph.D.
Sequencing by: Washington University Genome Sequencing Center
e distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LINL, send email to:
image.linl.gov
rimer: -40UP from Gibco
quality sequence stop: 175.

```
1. 196
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:322115"
```

```

/tissue_type="pooled germ cell tumors"
/lab_host="DHI10B"
/clone_lib="NCI_CGAP_GC6"
/polylinker="pRTT3-Pac
/site_1: Not I; Site 2: Eco RI; Plasmid
from the normalized library NCI_CGAP GC4 was pref
ss circles were made in vitro. Following HAP puri
this DNA was used as tracer in a subtractive hybr
reaction. The driver was PCR-amplified cDNAs from
of 5,000 clones made from the same library (clone
1257096-1258631, 1469064-1470983, and 1475592-147
Subtraction by Bento Soares and M. Fatima Ronald

```

ORIGIN	Query Match	12.3%	Score 169;	DB 10;	Length 196;
	Best Local Similarity	100.0%;	Pred. No. 1.4e-74;		
	Matches 169;	Conservative 0;	Mismatches 0;	Indels 0;	G
QY	1205	TAAGAGGGCTGGACCTGGCGGCGAGGAGGCCAAAGAGACTGGCGCTAGGCCAGGAC			
Db	190	TAAGAGGGCGCTGGACCTGGCGGCGAGGAGGCCAAAGAGACTGGCGCTAGGCCAGGAC			
QY	1265	CAAAATGTGAGGGGCGGAGAAACAAGAACAAGCTCCTCCCTTGAGAAATTCCTCTGTGGAT			
Db	130	CAAAATGTGAGGGGCGGAGAAACAAGAACAAGCTCCTCCCTTGAGAAATTCCTCTGTGGAT			
QY	1325	AAAAACAGATATATTTTTTATTTATTTATTTATTTGACAAAAATGTTGATAAATGG			1373
Db	70	AAACACGATATATTTTTTATTTATTTATTTATTTGACAAAAATGTTGATAAATGG			22

RESULT 68	BI677255/C	BI677255	422 bp	mRNA	linear	EST 1'
LOCUS		id87a02.x1	Human insulinoma	Homo sapiens	cdna 3'	mRNA seq
DEFINITION		BI677255				
ACCESSION		BI677255.1				
VERSION		GI:15630162				
KEYWORDS		EST.				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele-
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
	1 (bases 1 to 422)
REFERENCE	Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestn-
AUTHORS	Lemshinka,I., Searce,M., Brestelli,J., Gradwohl,G., Clift-
	Hillier,L., Marra,M., Page,D., Wylie,T., Martin,J., Blis-
	Schmitt,A., Theising,B., Ritter,E., Runko,I., Bennett,J.,
	Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvi
	Williams,T., Jackson,Y. and Bowers,Y.

TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Ino
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, C
MA 02138

Tel.: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. J. Ferrer. In vivo mass-exo
palescript SK- by Dr. H. Inoue. DNA sequencing by: Washin
University Genome Sequencing Center For information on ob
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.
Seq primer: -40UP from Gibco
High quality sequence stop: 416.

```

FEATURES
  source
    Location/Qualifiers
      1. 422
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /tissue_type="insulinoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="Human insulinoma"

```

/note="Organ: pancreas; Vector: pBluescript SK-; Site 1: XhoI; Site 2: EcoRI; Constructed with lambda ZAPII (Stratagene) by Dr. J. Ferrer, in vivo mass-excise pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

12.2%; Score 167; DB 12; Length 422;
 rity 100.0%; Pred. No. 1.7e-73; Indels 0; Gaps 0;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 TGCTGGGCGCTTTCACGTGTTTCCATCCACATAAATACAGTATCCCACTCT 1036
 TGCTGGGCGCTTTCACGTGTTTCCATCCACATAAATACAGTATCCCACTCT 363
 TACAACTCCCAACGGCCACTTCCACCTCACTAGCTCCCAATCCCTGACCT 1096
 TACAACTCCCAACGGCCACTTCCACCTCACTAGCTCCCAATCCCTGACCT 303
 TACAACTCCCAACGGCCACTTCCACCTCACTAGCTCCCAATCCCTGACCT 1143
 TACAACTCCCAACGGCCACTTCCACCTCACTAGCTCCCAATCCCTGACCT 256

16 372 bp mRNA linear EST 17-SEP-2001
 16 y1 Human insulinoma Homo sapiens cDNA 5', mRNA sequence.

16.1 GI:15630163

apiens (human)

apiens
 ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ses 1 to 372)
 D., Brown, J., Kent, G., Permutt, A., Lee, C., Kaestner, K.,
 ca, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 C.J., Marra, M., Page, D., Wylie, T., Martin, J., Blistain, A.,
 J.A., Theising, B., Ritter, S., Ronko, I., Bennett, J.,
 as, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,
 as, T., Jackson, Y., and Bowers, Y.
 ine Pancreas Consortium
 ished (2000)
 ESTs: id87a02.x1

1: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 ine Pancreas Consortium
 a University, Howard Hughes Medical Institute
 f Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 38

17-495-1812
 17-495-8557
 dmelton@biohph.harvard.edu
 y was constructed by Dr. J. Ferrer In vivo mass-excised to
 cript SK- by Dr. H. Inoue DNA sequencing by: Washington
 sity Genome Sequencing Center For information on obtaining a
 please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
 mer: -40RP from Gibco.

Location/Qualifiers
 1..372
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="insulinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Human insulinoma"
 /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:

XhoI; Site 2: EcoRI; Constructed with lambda ZAPII (Stratagene) by Dr. J. Ferrer, in vivo mass-excise pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

ORIGIN

Query Match 11.8%; Score 162; DB 12; Length 372;
 Best Local Similarity 100.0%; Pred. No. 6e-71;
 Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1145 GCATTGTGTTCACTGTTCTGTGGCGCAGGATGGTCCAGAGACCCCACTTCAG
 DB 211 GCATTGTGTTCACTGTTCTGTGGCGCAGGATGGTCCAGAGACCCCACTTCAG
 QY 1205 TAAGAGGGGCTGGACCTGGCGCGCAGGAGCCAAAGAGACTGGGCGCTAGGCCAGGAG
 DB 271 TAAGAGGGGCTGGACCTGGCGCGCAGGAGCCAAAGAGACTGGGCGCTAGGCCAGGAG
 QY 1265 CAAATGTAGGGGGCGAGAGAACAGACAGCTCTCCCTTGAG 1306
 DB 331 CAAATGTAGGGGGCGAGAGAACAGACAGCTCTCCCTTGAG 372

RESULT 70

AQ890280/c
 LOCUS
 DEFINITION
 HS 3188 B1 F05 MR CIT Approved Human Genomic Sperm Library
 sapiens genomic clone Plate=3188 Col=9 Row=L, genomic survey
 sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 1 (bases 1 to 436)
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzm
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams
 Hood, L.

TITLE
 Sequence-tagged connectors: A sequence approach to mapping
 scanning the human genome

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887

Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resge
 BAC end Web Server: http://www.htsc.washington.edu
 Plate: 3188 Row: L column: 9
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 436.

FEATURES

Location/Qualifiers
 1..436
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="plate=3188 Col=9 Row=L"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Libr
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clor

E-Coli DH10B"

11.8%; Score 162; DB 28; Length 436;

arity 99.0%; Pred. No. 6.2e-71;

conservative 0; Mismatches 3; Indels 0; Gaps 0;

CCTGTTACAGTGTTCCTCATCCACATAAATACAGTATTCCTCTTATCTTACA 1045

|||||

CCTGTTACAGTGTTCCTCATCCACATAAATACAGTATTCCTCTTATCTTACA 261

|||||

CCCCACCCCTCTCTCCACCTCTACTAGCTCCCAATCCCTGACCCCTTTGAGGCC 1105

|||||

CCCCACCCCTCTCTCCACCTCTACTAGCTCCCAATCCCTGACCCCTTTGAGGCC 201

|||||

TGATCTCACTCCCTCTGCGCAGACGCCCTGAGGCGATTTGTCTACTGTACTCT 1165

|||||

TGATCTCACTCCCTCTGCGCAGACGCCCTGAGGCGATTTGTCTACTGTACTCT 141

|||||

KCAAGGATGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGCTGACCTGGCG 1225

|||||

KCAAGGATGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGCTGACCTGGCG 81

|||||

KGAAGCAAGAGACTGGGCTAGGCGGCTAGGCGGCTGACCTGGCGGCGAGAAC 1285

|||||

KGAAGCAAGAGACTGGGCTAGGCGGCTAGGCGGCTGACCTGGCGGCGAGAAC 21

|||||

KCAAGCTCTCTCC 1300

|||||

KCAAGCTCTCTCC 6

|||||

279 409 bp mRNA linear EST 27-OCT-1999

14.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2622510 3',

sequence.

279

279.1 GI:6132886

sapiens (human)

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ases 1 to 409)

3AP http://www.ncbi.nlm.nih.gov/ncicgap.

nal Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Gene Index

lished (1997)

ct: Robert Strausberg, Ph.D.

: cgapbs-r@mail.nih.gov

a Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

t-Buck, M.D., Ph.D.

Library Preparation: Life Technologies, Inc.

Library Arrayed by: Greg Lennon, Ph.D.

Sequencing by: Washington University Genome Sequencing Center

e distribution: NCI-CGAP clone distribution information can be

through the I.M.A.G.E. Consortium/LLNL at:

io.llnl.gov/bbrp/image/image.html

primer: -40UP from Gibco

quality sequence stop: 405.

Location/Qualifiers

1..409

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2622510"

/tissue_type="poorly differentiated adenocarcinoma with

signed ring cell features"

/lab_host="DH10B"

/clone_lib="NCI_CGAP Gas4"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

ORIGIN

Query Match 11.5%; Score 158; DB 10; Length 409;
Best Local Similarity 100.0%; Pred. No. 6.8e-69;
Matches 158; Conservative 0; Mismatches 0; Indels 0; G

QY 1216 GGACTGGCGCAGAGCCAAAGAGAGCTGGCCTAGGCCAGGAGTCCCAATGT

|||||

DB 184 GGACTGGCGCAGAGCCAAAGAGAGCTGGCCTAGGCCAGGAGTCCCAATGT

|||||

QY 1276 GCGGAGAAAAGAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTTAAACAG

|||||

DB 124 GCGGAGAAAAGAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTTAAACAG

|||||

QY 1336 TATTTTATTTATTTATTTGACAAAATGTTGATAAATGG 1373

|||||

DB 64 TATTTTATTTATTTATTTGACAAAATGTTGATAAATGG 27

|||||

RESULT 72

R55379

LOCUS R55379 345 bp mRNA linear EST 2;

DEFINITION YJ77a08.r1 Soares breast 2NBH8t Homo sapiens cDNA clone

IMAGE:154742 5', mRNA sequence.

ACCESSION R55379

VERSION R55379.1 GI:824674

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 345)

AUTHORS Hillier, L., Clark, N., Dubucque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra,

Parsons, J., Rittkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P., et

Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 659

High quality sequence stops: 235 Source: IMAGE Consortium

This clone is available royalty-free through LLNL; conta

IMAGE Consortium (info@image.llnl.gov) for further inform

Insert Length: 659 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 235.

Location/Qualifiers

1..345

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:556839"

/db_xref="taxon:9606"

/clone="IMAGE:154742"

/sex="Female"

/dev_stage="adult"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares breast 2NBH8t"

/note="Organ: breast; Vector: p77T3D (Pharmacia)

modified polylinker; Site 1: Not I; Site 2: Eco

strand cDNA was primed with a Not I - oligo(dt)

TGTTACCAATCTAAGTGGAGCGGCCCTTTTITTTTTTTT

double-stranded cDNA was ligated to Eco RI adapt

(Pharmacia), digested with Not I and cloned into

and Eco RI sites of a modified p77T3 vector (Pha

Library went through one round of normalization

from normal prostate bulk tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo. "

9.5%; Score 130; DB 9; Length 179;
arity 100.0%; Pred. No. 1.1e-54; Indels 0; Gaps 0;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
CCTAGGCCAGGAGTCCCAATGTGAGGGCGAGAAACAAGACAAAGTCTCCCTT 1303
CCTAGGCCAGGAGTCCCAATGTGAGGGCGAGAAACAAGACAAAGTCTCCCTT 112
ATTCCTGTGGATTTTAAACAGATATTTTATTATTATTATTATTATTATTATT 1363
ATTCCTGTGGATTTTAAACAGATATTTTATTATTATTATTATTATTATTATT 52

AAATGG 1373
|||||
AAATGG 42

61 210 bp mRNA linear EST 13-AUG-1998
2.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1632506 3',
sequence.

61.1 GI:3418453

sapiens (human)

sapiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

uses 1 to 210)

AP http://www.ncbi.nlm.nih.gov/ncicgap.

al Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Gene Index

ished (1997)

t: Robert Strausberg, Ph.D.

: cgabbs@mail.nih.gov

a Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

;-Buck, M.D., Ph.D.

Library Preparation: M. Bento Soares, Ph.D.

Sequencing by: Greg Lennon, Ph.D.

a distribution: NCI-CGAP clone distribution information can be
through the I.M.A.G.E. Consortium/LINL at:

io.llnl.gov/bbrp/image/image.html

clmer: -40ml3 fwd. ET from Amersham

quality sequence stop: 203.

Location/Qualifiers

1. .210

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1632506"

/sex="female, pooled"

/tissue_type="breast"

/lab host="DH10B"

/clone lib="NCI CGAP Br2"

/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled bulk
breast tumor tissue, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. This library is the normalized version of
NCI CGAP Br1.1. Library was constructed by Bento Soares
and M. Fatima Bonaldo. "

ORIGIN

Query Match 9.3%; Score 128; DB 9; Length 210;
Best Local Similarity 100.0%; Pred. No. 1.2e-53;
Matches 128; Conservative 0; Mismatches 0; Indels 0; G

Qy 1246 GGCTAGGCCAGGAGTCCCAATGTGAGGGCGAGAAACAAGACAAAGTCTCTCC
Db 150 GGCTAGGCCAGGAGTCCCAATGTGAGGGCGAGAAACAAGACAAAGTCTCTCC
Qy 1306 GAATTCCTGTGGATTTTAAACAGATATTTATTATTATTATTATTATTGACAAAA
Db 90 GAATTCCTGTGGATTTTAAACAGATATTTATTATTATTATTATTATTGACAAAA
Qy 1366 ATAAATGG 1373
Db 30 ATAAATGG 23

Search completed: April 8, 2004, 23:42:07
Job time : 3973 secs

6:25:14 2004

us-09-245-198a-3.oligo.rng

GenCore version 5.1.6
copyright (c) 1993 - 2004 Compugen Ltd.

c search, using sw model

il 8, 2004, 19:06:21 ; Search time 654 Seconds
(without alignments)
8918.618 Million cell updates/sec

09-245-198A-3

3
tgtcatgttagcttga.....gacaaaatgttgataaatgg 1373

GO_NUC

-op 60.0 , Gapext 60.0

3863 segs, 2124099041 residues

s satisfying chosen parameters: 6747726

h: 0

h: 2000000000

sting first 100 summaries

Geneseq_29Jan04:*

Geneseqn1980s:*

Geneseqn1990s:*

Geneseqn2000s:*

Geneseqn2001as:*

Geneseqn2001bs:*

Geneseqn2002as:*

Geneseqn2003bs:*

Geneseqn2003cs:*

Geneseqn2004s:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
ed by analysis of the total score distribution.

SUMMARIES

ch	Length	DB	ID	Description
0.0	1373	2	AAV18600	AAV18600 Homo sapi
3.6	1306	7	ACC57587	ACC57587 Polynucle
3.6	1306	7	ACC57901	ACC57901 Human TWE
3.6	1306	9	ADC35205	ADC35205 Human CDN
0.8	1364	6	ABK34881	ABK34881 Human CDN
5.4	1353	3	AAA49717	AAA49717 Human PRO
5.4	1353	6	ABK40255	ABK40255 cDNA enco
5.4	1421	2	RAA55000	RAA55000 Human tum
9.8	1236	2	AAV47613	AAV47613 TNF relat
9.8	1236	4	RAA04350	RAA04350 Human TRE
7.1	1030	2	RAA23424	RAA23424 Human TNR
5.5	898	4	RAA03964	RAA03964 Expressio
9.6	493	8	ACH34013	ACH34013 Human end
3.0	195	6	ABK29540	ABK29540 Colon ade
5.5	282	2	AAAT22190	AAAT22190 Human gen
4.7	1239	9	ADC97713	ADC97713 Murine FL
4.4	60	6	ABN58848	ABN58848 Human spl
4.4	60	6	ABN41049	ABN41049 Human spl
4.4	60	6	ABN58593	ABN58593 Human spl
4.4	60	6	ABN58591	ABN58591 Human spl
4.4	60	6	ABN58849	ABN58849 Human spl
3.6	50	2	AAAX56002	AAAX56002 Human tum
3.6	50	3	AAA49732	AAA49732 Human PRO

TCTCGACTCCCTCGGCCACAGACCCCCAGGGCAATTGTGTTCACCTGTACTCTGTG 1080
AGGATGGGTCCAGAAGACCCCACTTTCAGGCACCTAAGAGGGGCTCGCACTCGCGGCA 1228
AGGATGGGTCCAGAAGACCCCACTTTCAGGCACCTAAGAGGGGCTGGACCTCGCGGCA 1140
GCCAAGAGACTGGGCTTAGCGCAGGAGTTCCTCAAACTCTGAGGGGCGAGAACACAG 1288
GCCAAGAGACTGGGCTTAGCGCAGGAGTTCCTCAAACTGTGAGGGCGAGAACACAG 1200
GCTCCTCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATATATTTATTATT 1348
GCTCCTCCCTTGAGAAATCCCTGTGGATTTTAAACAGATATATATTTATTATT 1260
GTGACAAAATGTTGATAAATGG 1373
GTGACAAAATGTTGATAAATGG 1285

Standard: cDNA: 1306 BP.

(first entry)

coding sequence.

; tumour necrosis factor; ligand; cytostatic;
tor; osteopathic; gene: ss.

```
Location/Qualifiers
18. .767
/*tag= a
/product= "Human TWI
```

-A2.

2002WO-US023782

2001US-0307838P.

N GENOME SCI INC.

Rosen CA:

0659/40.
2315.

oligomeric complex having a first polypeptide member of the TNF family, a second polypeptide member of the TNF family, and a third polypeptide member of the TNF family, useful for treating cancer, osteoporosis or an autoimmune disease.

Page 367-368: 388pp: English.

sequence is that of a polynucleotide encoding human TNFAK. In relation to compositions comprising heterotrimeric complexes of TNF ligand family members, and their use in the prevention and treatment of disease. In one embodiment, the heterotrimeric complex comprises full-length or extracellular TNFAK and full-length or extracellular portions of other TNF family members, preferably VEGI or VEGI-SV. The heterotrimeric complex is useful for treating an autoimmune disease, cancer, or other disease, and is particularly useful for inhibiting cancer cell proliferation, or inducing apoptosis of cancer cells, increasing B cell proliferation, or inducing apoptosis of T cells.

Db	1295	TTCCCTGTGGATTTTAAACAGATATTATTTTATTATTATTTGTCACAAAATGTC
Qy	1369	AATGG 1373
Db	1355	AATGG 1359
RESULT 9		
AAV47613		
ID	AAV47613	standard; cDNA; 1236 BP.
XX	AAV47613;	
AC		
XX	27-OCT-1998	(first entry)
DT		
XX	TNF related endothelium proliferative agent gene.	
DE		
XX	ss; TNF; endothelium proliferative agent; TREPA; wound healing; c	
KW	tissue grafting; vascularisation; apoptosis; autoimmune; birth c	
KW		
XX	Homo sapiens.	
OS		
XX		
XX	Key	Location/Qualifiers
PH	CDS	1..750
FT		/*tag= a
FT		/product= "TREPA"
XX		
XX	WO9835061-A2.	
PN		
XX	13-AUG-1998.	
PD		
XX	12-FEB-1998;	98WO-US002859.
PF		
XX	12-FEB-1997;	97US-00798692.
PR	10-FEB-1998;	98US-00021706.
PR		
XX	(ABBO) ABBOTT LAB.	
PA		
XX	Wiley SR;	
PI		
XX		
XX	WPI; 1998-447255/38.	
DR	P-PSDB; AAW29745.	
DR		
XX	Detecting nucleic acid encoding TREPA - useful for diagnosis and	
PT	treatment of autoimmune disease, tumours and inflammation.	
PT		
XX	Claim 11; Page 123-4; 142pp; English.	
PS		
XX	The TNF-related endothelium proliferative agent (TREPA), or its	
CC	activators or agonists, are used to treat a deficit of TREPA, e	
CC	promote wound healing or tissue grafting, by promoting vascular	
CC	also to induce apoptosis for treating cancer and eliminating aut	
CC	T cells, as an adjunct to cancer chemotherapy or antiviral treat	
CC	TREPA peptides can also be used to target cytotoxic agents or fi	
CC	affinity isolation of the corresponding receptor, the nucleic ac	
CC	which can be used to transform tumour cells to render them more	
CC	responsive to TREPA and to screen for TREPA mimics. Ribozymes, i	
CC	kNA , antibodies or peptides, are used to treat TREPA-associated	
CC	diseases, e.g. tumours and metastases (by inhibiting vasculati	
CC	inflammation or a wide range of autoimmune conditions, condition	
CC	involving abnormal stimulation of epithelial cells (e.g.	
CC	atherosclerosis), for birth control (inhibiting ovulation and p	
CC	formation) or other angiogenic conditions (e.g. ulcers)	
XX		
XX	Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;	
SQ		
	Query Match	69.8%; Score 958; DB 2; Length 1236;
	Best Local Similarity	99.6%; Pred. No. 0;
	Matches 1208; Conservative	0; Mismatches 5; Indels 0;
Qy	129	GGCAGGGGGGCCGGGGGAGCGGGCAGCGCCCTGCTGCTCCGTCGTCGCTG

BP; 223 A; 317 C; 279 G; 211 T; 0 U; 0 Other;
57.1%; Score 784; DB 2; Length 1030;
arity 99.9%; Pred. No. 0;
nservative 0; Mismatches 1; Indels 0; Gaps 0;
TTTGGGGAGCCGGGCATCGTGTCCGCCACGAGCCTGCCACGAGGAGCTGGTG 288
TTTGGGAGCCGGGCATCGTGTCCGCCACGAGGAGCTGCCACGAGGAGCTGGTG 60
GGAGGACGAGGACCCGTCGGGAATCCCAAGAGAGAAAGCCAGGATCCT 348
GGAGGACGAGGACCCGTCGGGAATCCCAAGAGAGAAAGCCAGGATCCT 120
TTTCTGGAACCGACTAGTTGGCCCTCGCAGAAGTGCACCTAAAGGCGCGAAAACA 408
TTTCTGGAACCGACTAGTTGGCCCTCGCAGAAGTGCACCTAAAGGCGCGAAAACA 180
TCGAAGAGCGATCGCAGCCCATTTATGAAGTTTCATCCACGACCTGGGACAGACGGA 468
TCGAAGAGCGATCGCAGCCCATTTATGAAGTTTCATCCACGACCTGGGACAGACGGA 240
GGCAGGTGTGACGGGACAGTGAAGTGGCTGGGAGGAAGCCAGAAATCAACAGCTCC 528
GGCAGGTGTGACGGGACAGTGAAGTGGCTGGGAGGAAGCCAGAAATCAACAGCTCC 300
TCTGCGCTACACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 588
TCTGCGCTACACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTAC 360
GTACTGTCAAGTGCACCTTGATGAGGGGAAGGCTGTCTACTCAAGCTGGACTTG 648
GTACTGTCAAGTGCACCTTGATGAGGGGAAGGCTGTCTACTCAAGCTGGACTTG 420
TGGATGTGTGTGGCCCTGCGCTGCTGGAGGAATTTCTAGCCACTGGGCGCAGT 708
TGGATGTGTGTGGCCCTGCGCTGCTGGAGGAATTTCTAGCCACTGGGCGCAGT 480
TCGGGCCCCAGCTCCGGCTTCGCCAGGTGTCTGGGCTGTGGCCCTGCGGCGCAGG 768
TCGGGCCCCAGCTCCGGCTTCGCCAGGTGTCTGGGCTGTGGCCCTGCGGCGCAGG 540
TCTGCGGATCCGCACCCCTCCCTGGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 828
TCTGCGGATCCGCACCCCTCCCTGGGGCCCATCTCAAGGCTGCCCCCTTCCTCACC 600
TCGGACTCTTCCAGGTTCACTGAGGGGCCCTGGTCTGCCCAAGTGTGCCAGGCT 888
TCGGACTCTTCCAGGTTCACTGAGGGGGCCCTGGTCTGCCCAAGTGTGCCAGGCT 660
GCTCCCTCGACAGCTCTCTGGGACCCGGTCCCTCTGGCCCAACCCCTCAGCCGCT 948
GCTCCCTCGACAGCTCTCTGGGACCCGGTCCCTCTGGCCCAACCCCTCAGCCGCT 720
GCTCCAGACTGTGCCCTCCCTCTAGAGGCTGTGCTGGGCTGTTCACGTGTTTTCCA 1008
GCTCCAGACTGTGCCCTCCCTCTAGAGGCTGTGCTGGGCTGTTCACGTGTTTTCCA 780
ACATAAATACAGTATTCACACTCTTATCTTACAACTCCGCCCAACGCCCACT 1063
ACATAAATACAGTATTCACACTCTTATCTTACAACTCCGCCCAACGCCCACT 835

dard; DNA; 898 BP.

(first entry)

ctor pDC409-L2-TWEAK fusion protein-encoding DNA.

TCCTGAACCCGACTAGTTTCGGCTCGCAGAGTGCACCTAAAGGCGCGAAACACGG 411
TCCTGAACCCGACTAGTTTCGGCTCGCAGAGTGCACCTAAAGGCGCGAAACACGG 429
GAAGAGCGATCGCAGCCCAATTATGAAGTTTATCCACGACCTGGACAGGACGGAGCG 471
GAAGAGCGATCGCAGCCCAATTATGAAGTTTATCCACGACCTGGACAGGACGGAGCG 489
CAGGTGTGGACGGACAGTGAAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC 531
CAGGTGTGGACGGACAGTGAAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC 549
TGCCTACACCCGACGATCGGGAGTTTATAGTCACCCGGCTGGCTCTACTAC 591
TGCCTACACCCGACGATCGGGAGTTTATAGTCACCCGGCTGGCTCTACTAC 609
ACTGTCAAGGTGCACTTTGATGAGGGAGGCTGTCTACTGAACTGGACTTGTG 651
ACTGTCAAGGTGCACTTTGATGAGGGAGGCTGTCTACTGAACTGGACTTGTG 669
ATGTGTGTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 711
ATGTGTGTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 729
TGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 771
TGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 789
TGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 831
TGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 849
TGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 856
TGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 874
Standard; cDNA; 493 BP.
(first entry)
Alial cell cDNA #2146.
quencing by hybridisation; SBH; expressed sequence tag; EST;
ig; biodiversity; genetic disorder.
-A1.
2001US-00918995.
2001US-00918995.
NAC R. T.
T. I.
HE-CRAIN B.
SON M C.
S L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX New polynucleotide sequences obtained from various cDNA libraries
PT as hybridization probes, as oligomers for PCR, for chromosome and
PT mapping, in the recombinant production of protein, or in generati
PT antisense DNA or RNA.
XX Claim 1; SEQ ID NO 21225; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising an
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequ
CC determined by the technique of SBH (sequencing by hybridisation).
CC included is a purified polypeptide comprising a sequence correspo
CC a reading frame of the novel polynucleotide. The nucleic acid seq
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human
CC in forensics, in assessing biodiversity, or in identifying muta
CC responsible for genetic disorders and other traits. The nucleotid
CC sequences are also useful as hybridisation probes, as oligomers f
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified poly
CC is useful for generating antibodies specific for it. The present
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequen
CC for this patent did not form part of the printed specification, b
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX Sequence 493 BP; 87 A; 180 C; 120 G; 101 T; 0 U; 5 Other;
SQ Query Match 29.6%; Score 407; DB 8; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.6e-179;
Matches 407; Conservative 0; Mismatches 0; Indels 0; C
QY 874 CAGTCGTCCAGGCTGCGGCTCCCTCGACAGCTCTCTGGGACCCCGTCCCTC
DB 87 CAGTCGTCCAGGCTGCGGCTCCCTCGACAGCTCTCTGGGACCCCGTCCCTC
QY 934 CCACCTCAGCGCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCTGCTG
DB 147 CCACCTCAGCGCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCTGCTG
QY 994 TTCAGTGTTCATCCCATATAAATACAGTATTCCTTATCTTACAACT
DB 207 TTCAGTGTTCATCCCATATAAATACAGTATTCCTTATCTTACAACT
QY 1054 ACCGCCACTCTCCACTCAGTGTCCCAATCCCTGACCTTTGAGGCCCCCA
DB 267 ACCGCCACTCTCCACTCAGTGTCCCAATCCCTGACCTTTGAGGCCCCCA
QY 1114 CTCGACTCCCCCTGGCCACAGACCCCCCAGGGCAATGTGTCTACTGTACTCTGT
DB 327 CTCGACTCCCCCTGGCCACAGACCCCCCAGGGCAATGTGTCTACTGTACTCTGT
QY 1174 GGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGGGCA
DB 387 GGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGGGCA
QY 1234 CCAAGAGACTGGGCTAGGCGCAGAGTCCCAATGTAGGGGCGA 1280
DB 447 CCAAGAGACTGGGCTAGGCGCAGAGTCCCAATGTAGGGGCGA 493
RESULT 14
ABK29540
ID ABK29540 standard; cDNA; 195 BP.
XX AC
XX AC
XX AC
XX AC
XX 23-APR-2002 (first entry)
XX

cinoma-specific cDNA #66.

denocarcinoma; colon cancer; tumour; gene; ss.

001WO-US018574.

000US-0210667P.

000US-0252614P.

A CORP.

ting GE, Xu J, Secrist H;

52/13.

polynucleotide encoding a polypeptide comprising a portion
of protein, for detection, diagnosis and therapy of human

133; 211pp; English.

relates to an isolated polynucleotide (I) encoding a
(I) comprising a portion of a colon tumour protein. A new
the (III) that hybridizes to (I) is useful for determining
of a cancer in a patient. (II) or antigen presenting cells
are useful for stimulating and/or expanding T cells
(I) tumour protein, by contacting T cells with (I), (II) or
ring cells that express (I), (II), or antigen
is that express (II) are useful for treating colon cancer
by incubating CD4+ and/or CD8+ T cells isolated from a
(I), (II), or antigen presenting cells that express (II), so
proliferate, and administering to the patient an effective
proliferated T cells, thus inhibiting the development of a
patient. A new composition is useful for stimulating an
in a patient. (I) or (II) is useful in vaccines and
compositions for prevention and treatment of colon cancer
agnosis and monitoring of the cancers. (I), (II) or an
st (II) is useful for detection, diagnosis and/or therapy
l cancer. (I) is useful as a probe or primer for nucleic
acid, and in the design and preparation of ribozyme
inhibiting expression of (II) in tumour cells. ABK29475-
assent human colon adenocarcinoma-specific cDNA sequences of

3P; 49 A; 51 C; 58 G; 37 T; 0 U; 0 Other;

13.0%; Score 179; DB 6; Length 195;

arity 100.0%; Pred. No. 5.3e-73;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

3TGTTCACCTACTCTGTGGCAGAGGATGGTCCAGAGACCCCACTTCAGGCAC 1205

3TGTTCACCTACTCTGTGGCAGAGGATGGTCCAGAGACCCCACTTCAGGCAC 76

3GGGCTGGACCTGGCGCAGAGACCCAAAGAGACTGGGCTAGGCCAGGAGTTCCC 1265

3GGGCTGGACCTGGCGCAGAGACCCAAAGAGACTGGGCTAGGCCAGGAGTTCCC 136

3TGAGGGCGAGAAACAGACAGCTCCTCCCTTGAGAAATTCCTGTGGATTTTT 1324

3TGAGGGCGGAGAAACAGACAGCTCCTCCCTTGAGAAATTCCTGTGGATTTTT 195

3ard; cDNA to mRNA; 282 BP.

AC AAT22190;

XX

DT 27-AUG-1996 (first entry)

XX

DE Human gene signature HUMGS03761.

XX

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.

OS Homo sapiens.

XX

PN W09514772-A1.

XX

PD 01-JUN-1995.

XX

PF 11-NOV-1994; 94WO-JP001916.

XX

PR 12-NOV-1993; 93JP-00355504.

XX

PA (MATS/) MATSUBARA K.

XX

PA (OKUB/) OKUBO K.

XX

PI Matsubara K, Okubo K;

XX

WPI; 1995-206931/27.

XX

PT Single-stranded DNA for identifying gene signatures - isolated fr
directed human cDNA library that reflects relative abundance of c
mRNA in specific human tissues.

XX

PS Claim 1; Page 1067; 2245pp; Japanese.

XX

CC A single-stranded DNA (or its complementary strand or the corresp
-stranded DNA) which comprises one of the 7837 "GS" sequences giv
AAT19001-T26837 and which is able to hybridise to part of human g
DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences w
obtained from 3'-directed cDNA libraries prepared from various hu
tissues; synthesis of cDNA was initiated from the 3'-end of mRNA
poly(T) as the sole primer. Since the 3'-untranslated sequence i
to a particular mRNA species, almost all the 3'-oriented cDNAs by
CC with specific mRNAs. Each library is constructed so as to reflect
CC accurately the relative abundance of different mRNAs in the parti
CC tissue from which it was derived. The appearance frequency of a g
in a cDNA library can be determined (esp. using primers and probe
CC derived from the GS sequences) as a means of diagnosing abnormal
CC function or for recognising different cell types

SQ Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 0 U; 5 Other;

Query Match 5.5%; Score 76; DB 2; Length 282;

Best Local Similarity 100.0%; Pred. No. 5.4e-25;

Matches 76; Conservative 0; Mismatches 0; Indels 0; G

QY 1146 CATTGTGTTCACTGTACTCTGTGGCAAGGATGGTCCAGAGACCCCACTTCAGG

Db 36 CATTGTGTTCACTGTACTCTGTGGCAAGGATGGTCCAGAGACCCCACTTCAGG

QY 1206 AAGAGGGGCTGGACCT 1221

Db 96 AAGAGGGGCTGGACCT 111

RESULT 16

ADC97713

ID ADC97713 standard; cDNA; 1239 BP.

XX

AC ADC97713;

XX

DT 15-JAN-2004 (first entry)

XX

DE Murine FL-TWEAK coding sequence.

XX

EAK; TNF relatedness and weak ability to induce cell death;
ecrosis Factor; TWEAK; fibrosis; cardiac disease;
lung disease; kidney disease; skin disease;
le disease; adipose tissue disease;
nal tract disease; pancreatic disease;
organ disease; neural disease; cartilage disease;
connective tissue disease; cellular death; hepatotropic;
1; gastrointestinal; osteopathic; gene; ss.

Location/Qualifiers
1..750
/*tag= a
/product= "FL-TWEAK"

A2.

2003WO-US011350.
2002US-0371611P.
N INC.

kubowski A, Zheng T, Hahn K;
1256/78.
7712.

NEAK-related condition, e.g. liver, gastrointestinal, kidney,
atic, cartilage or neural tissue condition in a subject
ministering to the subject a TWEAK agonist or antagonist.

3Q ID NO 2; 120pp; English.

sequence is the coding sequence for murine transmembrane FL-
relatedness and weak ability to induce cell death, where TNF
rosis Factor). TWEAK is a member of the TNF family. TWEAK
antagonists are useful for treating a TWEAK-related
g. fibrosis; cardiac disease; liver disease; lung disease;
se; skin disease; skeletal muscle disease; adipose tissue
ointestinal tract disease; pancreatic disease; reproductive
; neural disease; cartilage disease; bone disease;
issue disease; cellular death; and a pathological condition
expressing a TWEAK receptor.

9 BP; 249 A; 386 C; 331 G; 273 T; 0 U; 0 Other;
4.7%; Score 64; DB 9; Length 1239;
larity 100.0%; Pred. No. 2e-19;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GCCGCCGTCGGAGCCAGAGCGGAGGGGGCGCGGGGAGCGGCGGCGGCGGCTG 165
GCCGCCGTCGGAGCCAGAGCGGAGGGGGCGCGGGGAGCGGCGGCGGCGGCTG 60
G 169
G 64

ndard; DNA; 60 BP.

(first entry)

d transcript detection oligonucleotide SEQ ID NO:31596.

Human; mouse; rat; splice transcript; detection; RNA transcript;
splice variant; transcriptome; oligonucleotide library; ss.
Homo sapiens.
WO200210449-A2.
07-FEB-2002.
20-JUL-2001; 2001WO-IB001903.
28-JUL-2000; 2000US-0221607P.
02-MAY-2001; 2001US-0287724P.
(COMP-) COMPUEN INC.
Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
WPI; 2002-257383/30.
New oligonucleotide libraries comprising oligonucleotides which
selectively hybridize to mRNAs transcribed from a transcription
genome, useful for detecting tissue-, pathology-, and development
specific genes.
Example 1; SEQ ID NO 31596; 47pp; English.
The present invention describes oligonucleotide libraries for det
messenger RNAs that populate a (sub-)transcriptome, where the (s
)transcriptome comprises messenger RNAs transcribed from multiple
transcription units that populate a genome. The library comprises
oligonucleotides, each capable of hybridising selectively to a se
messenger RNAs transcribed from a given transcription unit of the
which encodes one or more messenger RNA splice variants. The
oligonucleotide libraries are useful for detecting mRNAs from a
biological sample, in expression profiling studies, in qualitati
quantitatively characterising the corresponding transcriptome, a
detecting RNA transcripts and splice variants of human or animal
transcriptomes. The libraries may also be used as specialised mi
libraries to detect transcripts of a sub-transcriptome under a p
biological or pathological state, and so allowing the detection
- and pathology-specific genes such as those genes only expres
specific tissue under a specific pathological condition; to dete
developmental specific genes; and to detect RNA transcripts and
variants of a transcriptome of a patient suffering from a partic
disorder. ABN27253 to ABN59589 represent oligonucleotide sequen
rats, humans and mice, which are used in the exemplification of
present invention. N.B. The sequence data for this patent did no
part of the printed specification, but was obtained in electroni
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 60 BP; 13 A; 16 C; 17 G; 14 T; 0 U; 0 Other;
SQ
Query Match 4.4%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.7e-17; Indels 0;
Matches 60; Conservative 0; Mismatches 0; Indels 0;
QY 1145 GCATTGTGTTCACTGTTCTGTGGCAAGGATGGGTCCAGAGACCCCACTTCA
DB 1 GCATTGTGTTCACTGTTCTGTGGCAAGGATGGGTCCAGAGACCCCACTTCA
RESULT 18
ABN41049
ID ABN41049 standard; DNA; 60 BP.
XX
AC ABN41049;
XX
15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:137
KW Human; mouse; rat; splice transcript; detection; RNA transcript;

; transcriptome; oligonucleotide library; ss.

:001WO-IB001903.

!000US-0221607P.

001US-0287724P.

;EN INC.

usserman A, Mintz E, Mintz L, Faiqler S;

183/30.

otide libraries comprising oligonucleotides which hybridize to mRNAs transcribed from a transcription unit of a ... for detecting tissue-, pathology-, and developmental-

) ID NO 13797; 47pp; English.

vention describes oligonucleotide libraries for detecting : that populate a (sub-)transcriptome, where the (sub- : comprises messenger RNAs transcribed from multiple : units that populate a genome. The library comprises several : , each capable of hybridising selectively to a set of : ; transcribed from a given transcription unit of the genome, : one or more messenger RNA splice variants. The : le libraries are useful for detecting mRNAs from a : ple, in expression profiling studies, in qualitatively or : / characterising the corresponding transcriptome, and in : transcripts and splice variants of human or animal : . The libraries may also be used as specialised mini : detect transcripts of a sub-transcriptome under a particular : pathological state, and so allowing the detection of tissue : MY-specific genes such as those genes only expressed in : under a specific pathological condition; to detect : specific genes; and to detect RNA transcripts and splice : transcriptome of a patient suffering from a particular : /7252 to ARN5989 represent oligonucleotide sequences from : and mice, which are used in the exemplification of the : ion. N.B. The sequence data for this patent did not form : inted specification, but was obtained in electronic format : WIPO at ftp.wipo.int/pub/published_pct_sequences/

, 17 A; 15 C; 15 G; 13 T; 0 U; 0 Other;

4.4%; Score 60; DB 5; Length 60;
urity 100.0%; Pred. No. 1.7e-17;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
LATGTGAGGGGCGAGAAACAAGCTCTCCCTTGAGAAATTCCTGTGGATTT 1322
LATGTGAGGGGCGAGAAACAAGCTCTCCCTTGAGAAATTCCTGTGGATTT 60

yard; DNA: 60 BP.

(first entry)

transcript detection oligonucleotide SEQ ID NO:31341.

rat; splice transcript; detection; RNA transcript;

transcriptome; oligonucleotide library; ss.

XX
OS Homo sapiens.

AA
PN
WO200210449-A2.

PD 07-FEB-2002.

20-JUL-2001; 20C1WO-IB001903.

PR 28-JUL-2000; 2000US-0221607P.

PR 02-MAY-2001; 2001US-0287724P.

PA (COMP-) COMPUGEN INC.

PI Shoshan A, Wasserman A, Mintz E, Mintz L, Paigler S;

DR WPI; 2002-257383/30.

PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription u
PT genome, useful for detecting tissue-, pathology-, and development
PT specific genes.

PS Example 1: SEO ID NO 31341: 47pp: English;

The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively characterising the corresponding transcriptome, and detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini-libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of - and pathology-specific genes such as those genes only expressed in a specific tissue under a specific pathological condition, to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular disorder. ABN27253 to ABN59589 represent oligonucleotide sequences, humans and mice, which are used in the exemplification of the present invention. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form from WIPO at ftb.wipo.int/pub/published and is directed to WIPO at ftb.wipo.int/pub/published.

Sequence 60 BP: 12 A: 19 C: 10 G: 19 T: 0 U: 0 Other: 0

Query Match 4.4%; Score 60; DB 6; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.7e-17;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; G:
 QY 979 GCGTCTGGGCGCTGTTTCACGIGTTTTCCATCCCAATAACAGPATTTCCCACT.
 Db 1 GCGTCTGGGCGCTGTTTCACGIGTTTTCCATCCCAATAACAGPATTTCCCACT.

RESULT 20

ABNS8591

ID ABN58591 standard; DNA; 60 BP.

AC ABN58591;

15-JUL-2002 (first entry)

XX
DE Human apliced transcript detection oligonucleotide SEO ID NO:31333

XX Human: mouse: rat: splice transcript: detection: RNA transcript: KW

KW splice variant; transcriptome; oligonucleotide library; ss.
KW splice variant; transcriptome; oligonucleotide library; ss.

XX

98WO-US021407.
97US-0062037P.
97US-0065962P.
TECH INC.
Marsters SA, Pitti R;
982/24.
3- ligand (a tumor necrosis factor) homologue.
je 36; 74pp; English.
vention describes a human tumor necrosis factor (TNF) and
mologue designated Apo-3 ligand. Apo-3 ligand has
ivity. Apo-3 ligand can be used to induce apoptosis in
er cells, to induce NF-kappaB-dependent transcription and
/SAPK-dependent responses in mammalian cells. The present
esents an Apo-3 ligand probe, which is used in an example
ent invention
P; 10 A; 18 C; 13 G; 9 T; 0 U; 0 Other;
arity 3.6%; Score 50; DB 2; Length 50;
100.0%; Pred. No. 7.7e-13;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
CCCTCTCGCTACACCGCCAGATCGGGGAGTTTATAGTCACCCGG 576
CCCTCTCGCTACACCGCCAGATCGGGGAGTTTATAGTCACCCGG 50
dard; DNA; 50 BP.
(first entry)
DNA probe.
; antitumor; tumor; therapy; cytostatic; breast cancer;
r; renal cancer; colorectal cancer; uterine cancer;
er; lung cancer; bladder cancer;
us system cancer; melanoma; leukaemia; neoplasm; probe; ss.
2.
98WO-US028565.
98US-0113296P.
98WO-US005028.
99US-0130232P.
99US-0131445P.
99US-0134287P.
99US-0144758P.
99US-0145698P.
98WO-US021090.
98WO-US021547.

PA (GETH) GENENTECH INC.
XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
PI Napier MA, Pitti RM, Wood WI;
XX WPI; 2000-442668/38.
XX Novel composition to inhibit neoplastic cell growth or for treati
PT in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866.
XX Example 3; Page 98; 172pp; English.
XX The present sequence is that of a DNA probe based on an isolated
CC expressed sequence tag showing homology to human Apo-2 ligand. Th
CC was used to screen a human foetal kidney cDNA library to identify
CC clone DNA30879-1152 (see AAA49717), which encodes human antitumou
CC protein PRO207 (see AAY95338). A claimed method for inhibiting th
CC of a tumour cell comprises exposing the tumor cell to PRO179, PRO
CC PRO320, PRO219, PRO224, PRO328, PRO301, PRO526, PRO362, P
CC PRO509 or PRO866 (see AAY95337-49). The tumour is especially brea
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder or c
CC nervous system cancer, melanoma or leukaemia. Nucleic acids encod
CC PRO179 etc. are used in the recombinant production of antitumour
CC polypeptides
XX Sequence 50 BP; 10 A; 18 C; 13 G; 9 T; 0 U; 0 Other;
SQ Query Match 3.6%; Score 50; DB 3; Length 50;
Best Local Similarity 100.0%; Pred. No. 7.7e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; G
QY 527 CCAGCCCTCTCGCTACACCGCCAGATCGGGGAGTTTATAGTCACCCGG 576
DB 1 CCAGCCCTCTCGCTACACCGCCAGATCGGGGAGTTTATAGTCACCCGG 50
RESULT 24
ABK40292
ID ABK40292 standard; DNA; 50 BP.
XX AC ABK40292;
XX 15-JUL-2002 (first entry)
DT Oligonucleotide probe for human PRO207 DNA.
DE Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
XX leukaemia; neuronal disorder; stromal disorder; blastocoele disc
KW inflammatory disorder; immune disorder; angiogenic disorder; cyt
KW neuroprotective; probe; ss.
XX Homo sapiens.
OS WO200153486-A1.
XX 26-JUL-2001.
XX 11-FEB-2000; 2000WO-US003565.
XX 08-MAR-1999; 99WO-US005028.
PR 11-MAR-1999; 99US-0123972P.
PR 11-MAY-1999; 99US-0133459P.
PR 02-JUN-1999; 99WO-US012252.
PR 22-JUN-1999; 99US-0140650P.
PR 22-JUN-1999; 99US-0140653P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 17-AUG-1999; 99US-0149395P.
PR 31-AUG-1999; 99US-0151689P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.

99WO-US028313.
99WO-US028301.
99WO-US028634.
2000WO-US000219.

TECH INC.

Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
Wood WI;

567/26.

nucleic acids encoding PRO polypeptides, useful for treating
ignant tumors, leukemias and lymphoid malignancies,
angiogenic and immunologic disorders.

ige 109; 302pp; English.

vention relates to the isolation of novel human PRO
(AAU86128-AAU86162) and the polynucleotide sequences
1. The PRO polypeptides, agonists, antagonists or anti-PRO
e useful for treating benign or malignant tumors (e.g.
, bladder, breast, etc), leukemias and lymphoid
other disorders such as neuronal, glial, astrocytal,
glandular, macrophagal, stromal and blastocoeleic disorders,
immune and angiogenic disorders. The polynucleotide
; also useful in gene therapy. The present sequence
probe used in the methods of the present invention

IP; 10 A; 18 C; 13 G; 9 T; 0 U; 0 Other;

3.6%; Score 50; DB 6; Length 50;

arity 100.0%; Pred. No. 7.7e-13; Indels 0; Gaps 0;
Conservative 0; Mismatches 0;

CCCTCTGGGCTACACCGCCAGATCGGGAGTTTATAGTCACCCGG 576
CCCTCTGGGCTACACCGCCAGATCGGGAGTTTATAGTCACCCGG 50

andard; DNA; 50 BP.

(first entry)

te gene expression profiling probe SEQ ID NO 2558.

; gene expression profiling; allograft rejection;
sis; congestive heart failure; systemic lupus erythematosus;
thrititis; osteoarthritis; cytomegalovirus; infection; probe;

A2.

2001WO-US047856.

2000US-0241994P.
2001US-0296764P.

ARDIA INC.

, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
ard R, Quertemous T, Johnson F;

5525/68.

XX New system for leukocyte expression profiling, diagnosing a disease
PT monitoring (the rate of) progression of a disease, e.g. atheroscl
PT or congestive heart failure, comprises diagnostic oligonucleotide
XX
XX Claim 1; Page 408; Opp; English.

XX The invention relates to a system for detecting gene expression,
CC comprises one or two isolated DNA molecules that detect expressi
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disease
CC predicting therapeutic outcome, determining prognosis for a patient
CC predicting disease complications in an individual or monitoring
CC to treatment in an individual. The diseases include cardiac allog
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythem
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ Sequence 50 BP; 14 A; 11 C; 19 G; 6 T; 0 U; 0 Other;

Query Match 3.6%; Score 50; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 7.7e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; C

OY 1196 TTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGAGGAGCAAGAGACTG 1245
DB 1 TTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGAGGAGCAAGAGACTG 50

RESULT 26

AAAX23425

ID AAX23425 standard; DNA; 701 BP.

XX
AC AAX23425;

DT 18-JUN-1999 (first entry)

DE Mouse TNRL3 DNA.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF
KW developmental abnormality; gestational abnormality; prostate c
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
KW cytoplasmic domain; immunogen; antibody preparation; breast carc
KW apoptosis; mouse; ss.

OS Mus sp.

Key Location/Qualifiers
CDS 1..636
FT /*tag= a
FT /product= "TNRL3"

XX WO9911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US018393.

XX 05-SEP-1997; 97US-00924634.

XX (UNIW) UNIV WASHINGTON.

XX Chaudhary PM;

XX WPI; 1999-205191/17.

XX P-PSDB; AAW93591.

XX New Tumor Necrosis Factor family receptor polypeptides and ligan
PT useful for diagnosis and treatment of prostate cancer and develo
PT or gestational abnormalities.

XX

Fig 13B; 156pp; English.

n describes isolated Tumor Necrosis Factor (TNF) family peptides: APO4, APO6, APO8 and APO9 or their active fragments. APO4 is useful for diagnosing prostate cancer by levels of APO4 in an individual. Prostate cancer can also be APO4 selective binding agents linked to a therapeutic polypeptides are also useful for identifying selective s, useful in diagnosis/treatment of disease by binding of polypeptide/active fragment which is extracellular, or the cell surface. The binding is preferably performed in lypeptides/ active fragments are also useful for screening and antagonists by binding and observing the change in APO4 active pharmacological agents useful in diagnosis or disease are also identified using APO4 polypeptides/active APO4 signal transducer molecules that specifically interact asmc domain of APO4 and detecting a change in level of APO4 method is performed in vivo or in vitro. APO polypeptides l as immunogens for preparing antibodies. APO4 is also agnosis/treatment of developmental or gestational . APO8 was transfected to human breast carcinoma cell line duced apoptosis

BP; 139 A; 210 C; 203 G; 149 T; 0 U; 0 Other;

3.4%; Score 46; DB 2; Length 701;

arity 100.0%; Pred. No. 5e-11; Indels 0; Gaps 0; conservative 0; Mismatches 0;

TGGGCTCTACTGCTACTGCTAGGTGACCTTTGATGAGG 620

TGGGCTCTACTGCTACTGCTAGGTGACCTTTGATGAGG 401

dard; cDNA; 1168 BP.

(first entry)

tumour necrosis factor related ligand (TRELL) gene.

necrosis factor related ligand; tnfr; treatment; cancer; sease; immune system; stimulation; suppression; on; ds.

Location/Qualifiers

2..679

/tag= a

/note= "tumour necrosis factor related ligand"

97WO-US013945.

96US-0023541P.

96US-0028515P.

97US-0040820P.

N INC.

GENEVA FACULTY MEDICINE.

e Y, Browning JL;

619/13.

524.

XX Tumour necrosis factor related ligand - useful for, e.g. treating
PT auto-immune disease and immune responses to tissue grafts.

XX Claim 2; Page 45-46; 69pp; English.

XX The sequence is that encoding mouse tumour necrosis factor relate
(TRELL). TRELL or active fragments can be included with a carrier
pharmaceutical compositions to treat cancer, autoimmune diseases
immune responses to tissue grafts, or to stimulate or suppress tl
system. It is useful to screen for TRELL receptors, by labelling
detectable label and screening compositions for binding. Agents
interfering with TRELL-receptor binding can also be screened for,
then be administered, optionally with interferon- gamma, to induc
death or treat, suppress or alter immune responses (especially in
human adenocarcinoma cells) involving a signal pathway between TI
its receptor. The DNA sequence can be used in gene therapy for TI
related disorders in mammals (especially humans), e.g. tumours,
autoimmune and inflammatory diseases or inherited genetic disorde
introducing into cells, and expressing, therapeutically effective
of a vector, e.g. a virus comprising a gene encoding TRELL. It m
be of use in the preparation of prepare probes for screening
natural/synthetic DNAs for TRELL-encoding sequences and for anti
therapy

XX Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 U; 0 Other;

Query Match 3.4%; Score 46; DB 2; Length 1168;

Best Local Similarity 100.0%; Pred. No. 4.9e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; C

OY 575 GGGCTGGGCTCTACTGCTAGGTGACCTTTGATGAGG 620

DB 399 GGGCTGGGCTCTACTGCTAGGTGACCTTTGATGAGG 444

RESULT 28

ABX37032

ID ABX37032 standard; cDNA; 408 BP.

AC ABX37032;

DT 20-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #2197.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identific
KW gene analysis; cattle breeding.

OS Bos Taurus.

PN US2002137139-A1.

XX 26-SEP-2002.

PF 24-SEP-2001; 2001US-00960352.

XX 12-JAN-1999; 99US-0115707P.

PR 11-JAN-2000; 2000US-00480902.

XX (BYATT// BYATT J C.

PA (MATH// MATHIALAGAN N.

PA (TAON// TAO N.

PA (WARR// WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat

PT deposition, useful for genome mapping, gene identification and ar
PT cattle breeding, or for genetically improving cattle.

99WO-US028301.
99WO-US028634.
2000WO-US000219.

TECH INC.

Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
Wood WI;

567/26.

ucleic acids encoding PRO polypeptides, useful for treating
ignant tumors, leukemias and lymphoid malignancies,
angiogenic and immunologic disorders.

age 140; 302pp; English.

nvention relates to the isolation of novel human PRO
(AAU86128-AAU86162) and the polynucleotide sequences
. The PRO polypeptides, agonists, antagonists or anti-PRO
e useful for treating benign or malignant tumours (e.g.
, bladder, breast, etc), leukaemias and lymphoid
other disorders such as neuronal, glial, astrocytal,
glandular, macrophagal, stromal and blastocoeic disorders,
immune and angiogenic disorders. The polynucleotide
also useful in gene therapy. The present sequence
PCR primer used in the methods of the present invention

P; 7 A; 1 C; 9 G; 9 T; 0 U; 0 Other;

arity 1.9%; Score 26; DB 6; Length 26;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

TCCACATAAATACAGTATCC 1030

TCCACATAAATACAGTATCC 1

dard; DNA; 140 BP.

(first entry)

ssion-related sequence, SEQ ID 542.

irucide; apoptotic; gene therapy; tumour suppression;
ion; apoptosis; virus resistance; viral infection; tumour;
tive disease; ds.

2.

2002WO-FR000543.

2001FR-00001925.

ULAR ENGINES LAB.

Amson R, Tuijnder M, Susini L;

286/05.

cid encoding a translationally controlled tumor protein,
eating, preventing and diagnosing viral, tumor or
diseases.

XX

PS Disclosure; Page: 45pp; French.

XX

CC The present invention relates to novel nucleic acid sequences (AE
CC ABZ79313), which are involved in the molecular pathways of tumour
CC suppression, tumour reversion, apoptosis and/or virus resistance.
CC sequences are also useful for treatment or prevention of viral, t
CC and cell degenerative diseases, and also for diagnosis and progn
CC these diseases. Note: The sequence data for this patent is not
CC represented in the printed specification but is based on sequence
CC information supplied by the European Patent Office

XX Sequence 140 BP; 43 A; 23 C; 28 G; 46 T; 0 U; 0 Other;

Query Match 1.9%; Score 26; DB 7; Length 140;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 26; Conservative 0; Mismatches 0; Indels 0; G

Qy 1 ATGTCATTGTAGACTTTGAAATTC 26

Db 67 ATGTCATTGTAGACTTTGAAATTC 42

RESULT 32

ABZ09382/c

ID ABZ09382 standard; DNA; 140 BP.

XX

AC ABZ09382;

DT 16-JAN-2003 (first entry)

XX

DE Human oligonucleotide SEQ ID 542.

XX

KW Human: tumour suppressor; virucide; cytostatic; nootropic;
KW neuroprotective; neuroleptic; gene therapy; tumour suppression;
KW tumour reversion; apoptosis; viral resistance; viral infection;
KW cell degeneration; Alzheimer's disease; schizophrenia; cancer; ds

XX Homo sapiens.

XX FR2822475-A1.

XX 27-SEP-2002.

XX 20-MAR-2002; 2002FR-00003459.

XX 13-FEB-2001; 2001FR-00001925.

XX

PA (MOLE-) MOLECULAR ENGINES LAB SA.

XX

PI Telerman A, Amson R, Tuijnder M, Susini L;

XX

DR WPI; 2003-032204/03.

XX

PT New human nucleic acid, useful for diagnosis, prognosis and treat
PT e.g. of tumors, also related vectors, transformed cell, polypepti
PT antibodies.

XX

PS Disclosure; Page 120; 189pp; French.

XX

CC The present invention relates to human oligonucleotides (ABZ08841
CC ABZ09860). The expression of the oligonucleotides is implicated i
CC suppression or reversion, apoptosis and/or viral resistance. The
CC oligonucleotides are useful for preventing and/or treating viral
CC infection, tumour development and cell degeneration (e.g. Alzheim
CC disease and schizophrenia), especially cancer

XX

XX Sequence 140 BP; 43 A; 23 C; 28 G; 46 T; 0 U; 0 Other;

Query Match 1.9%; Score 26; DB 7; Length 140;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 26; Conservative 0; Mismatches 0; Indels 0; G

06:25:14 2004

us-09-245-198a-3.oligo.rng

TCATTGTTAGACTTTGAAATTC 26
|||||
TCATTGTTAGACTTTGAAATTC 42

ndard; DNA; 145 BP.

(first entry)

ession-related sequence, SEQ ID 15.

virucide; apoptotic; gene therapy; tumour suppression;
sion; apoptosis; virus resistance; viral infection; tumour;
active disease; ds.

A2.

2002WO-FR000543.

2001FR-00001925.

CULAR ENGINES LAB.

Amson R, Tuijnder M, Susini L;

8286/05.

acid encoding a translationally controlled tumor protein,
reating, preventing and diagnosing viral, tumor or
diseases.

Page; 45pp; French.

invention relates to novel nucleic acid sequences (ABZ78294-
hich are involved in the molecular pathways of tumour
tumour reversion, apoptosis and/or virus resistance. The
e also useful for treatment or prevention of viral, tumour
enerative diseases, and also for diagnosis and prognosis of
es. Note: The sequence data for this patent is not
in the printed specification but is based on sequence
supplied by the European Patent Office

BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;

1.9%; Score 26; DB 7; Length 145;
larity 100.0%; Pred. No. 0.11;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TCATTGTTAGACTTTGAAATTC 26
|||||
TCATTGTTAGACTTTGAAATTC 99

ndard; DNA; 145 BP.

(first entry)

ession-related sequence, SEQ ID 135.

virucide; apoptotic; gene therapy; tumour suppression;
sion; apoptosis; virus resistance; viral infection; tumour;

KW cell degenerative disease; ds.

XX Unidentified.

OS WO200264731-A2.

PN 22-AUG-2002.

XX 13-FEB-2002; 2002WO-FR000543.

XX 13-FEB-2001; 2001FR-00001925.

PR (MOLE-) MOLECULAR ENGINES LAB.

XX Telerman A, Amson R, Tuijnder M, Susini L;

XX WPI; 2003-058286/05.

DR New nucleic acid encoding a translationally controlled tumor pro
PT useful for treating, preventing and diagnosing viral, tumor or
PT degenerative diseases.

XX Disclosure; Page; 45pp; French.

XX The present invention relates to novel nucleic acid sequences (A
CC ABZ79313), which are involved in the molecular pathways of tumou
CC suppression, tumour reversion, apoptosis and/or virus resistance
CC sequences are also useful for treatment or prevention of viral,
CC and cell degenerative diseases, and also for diagnosis and progn
CC these diseases. Note: The sequence data for this patent is not
CC represented in the printed specification but is based on sequenc
CC information supplied by the European Patent Office

XX Sequence 145 BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;

Query Match 1.9%; Score 26; DB 7; Length 145;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
|||||

Db 74 ATGTCATTGTTAGACTTTGAAATTC 99
|||||

RESULT 35

ABZ08955

ID ABZ08855 standard; DNA; 145 BP.

XX AC ABZ08855;

XX 16-JAN-2003 (first entry)

XX Human oligonucleotide SEQ ID 15.

XX Human; tumour suppressor; virucide; cytostatic; nootropic;
KW neuroprotective; neuroleptic; gene therapy; tumour suppression;
KW tumour reversion; apoptosis; viral resistance; viral infection;
KW cell degeneration; Alzheimer's disease; schizophrenia; cancer; d

OS Homo sapiens.

XX FR2822475-A1.

XX 27-SEP-2002.

XX 20-MAR-2002; 2002FR-00003459.

XX 13-FEB-2001; 2001FR-00001925.

XX (MOLE-) MOLECULAR ENGINES LAB SA.

XX Telerman A, Amson R, Tuijnder M, Susini L;

XX

204/03.
 leic acid, useful for diagnosis, prognosis and treatment.
 3, also related vectors, transformed cell, polypeptides and
 age 40; 189pp; French.
 nvention relates to human oligonucleotides (ABZ08841-
 e expression of the oligonucleotides is implicated in tumour
 r reversion, apoptosis and/or viral resistance. The
 des are useful for preventing and/or treating viral
 our development and cell degeneration (e.g. Alzheimer's
 chizophrenia), especially cancer
 BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;
 1.9%; Score 26; DB 7; Length 145;
 arity 100.0%; Pred.No. 0.11;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 CATTGTTAGACTTTGAAATTC 26
 |||||
 CATTGTTAGACTTTGAAATTC 99
 dard; DNA; 145 BP.
 (first entry)
 cleotide SEQ ID 135.
 suppressor; virucide; cytostatic; nootropic;
 ve; neuroleptic; gene therapy; tumour suppression;
 ion; apoptosis; viral resistance; viral infection;
 tion; Alzheimer's disease; schizophrenia; cancer; ds.
 2002FR-00003459.
 2001FR-00001925.
 JLAR ENGINES LAB SA.
 Amson R, Tuijnder M, Susini L;
 204/03.
 leic acid, useful for diagnosis, prognosis and treatment.
 3, also related vectors, transformed cell, polypeptides and
 age 58; 189pp; French.
 nvention relates to human oligonucleotides (ABZ08841-
 e expression of the oligonucleotides is implicated in tumour
 r reversion, apoptosis and/or viral resistance. The
 des are useful for preventing and/or treating viral
 our development and cell degeneration (e.g. Alzheimer's
 chizophrenia), especially cancer
 BP; 51 A; 28 C; 23 G; 43 T; 0 U; 0 Other;
 1.9%; Score 26; DB 7; Length 145;

Best Local Similarity 100.0%; Pred.No. 0.11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; G
 QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
 |||||
 Db 74 ATGTCATTGTTAGACTTTGAAATTC 99
 RESULT 37
 AAT19717
 ID AAT19717 standard; cDNA to mRNA; 147 BP.
 XX
 AC AAT19717;
 XX
 DT 05-JUL-1996 (first entry)
 XX
 DE Human gene signature HUMGS00791.
 XX
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequenc
 KW human; cloning; mapping; non-biased library; diagnosis; detection
 KW cell typing; abnormal cell function; ss.
 XX Homo sapiens.
 OS
 XX WO9514772-A1.
 PN
 XX 01-JUN-1995.
 PD
 XX 11-NOV-1994; 94WO-JP001916.
 PF
 XX 12-NOV-1993; 93JP-00355504.
 PR
 XX (MATS/) MATSUBARA K.
 PA (OKUB/) OKUBO K.
 XX
 XX Matsubara K, Okubo K;
 PI
 XX WPI; 1995-206931/27.
 DR
 XX
 XX Single-stranded DNA for identifying gene signatures - isolated fr
 PT directed human cDNA library that reflects relative abundance of c
 PT mRNA in specific human tissues.
 XX
 PS Claim 1; Page 456; 2245pp; Japanese.
 CC
 CC A single-stranded DNA (or its complementary strand or the corresp
 CC -stranded DNA) which comprises one of the 7837 "GS" sequences giv
 CC AAT19001-R26837 and which is able to hybridise to part of human g
 CC DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences w
 CC obtained from 3'-directed cDNA libraries prepared from various hu
 CC tissues; synthesis of cDNA was initiated from the 3'-end of mRNA
 CC poly(T) as the sole primer. Since the 3'- untranslated sequence i
 CC to a particular mRNA species, almost all the 3'-oriented cDNAs hy
 CC with specific mRNAs. Each library is constructed so as to reflect
 CC accurately the relative abundance of different mRNAs in the parti
 CC tissue from which it was derived. The appearance frequency of a g
 CC in a cDNA library can be determined (esp. using primers and probe
 CC derived from the GS sequences) as a means of diagnosing abnormal
 CC function or for recognising different cell types
 XX
 SQ Sequence 147 BP; 51 A; 28 C; 25 G; 43 T; 0 U; 0 Other;
 Query Match 1.9%; Score 26; DB 2; Length 147;
 Best Local Similarity 100.0%; Pred.No. 0.11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; G
 QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
 |||||
 Db 74 ATGTCATTGTTAGACTTTGAAATTC 99
 RESULT 38
 AAI69032/c

Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green,
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 218485)
Green, E.D.
Direct Submission
Submitted (17-JUN-2002) NIH Intramural Sequencing Center,
Gromvont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 218485)
Green, E.D.
Direct Submission
Submitted (05-JUN-2003) NIH Intramural Sequencing Center,
Gromvont Circle, Gaithersburg, MD 20877, USA
On Jun 5, 2003 this sequence version replaced gi:26449071
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoonhgri.nih.gov
----- Project Information
Center project name: cms
Center clone name: 145D13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indication order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g. human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.90319
Consensus quality: 214085 bases at least Q40
Consensus quality: 215255 bases at least Q30
Consensus quality: 216264 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 216885; sum-of-contents
Quality coverage: 12.65x in Q20 bases; agarose-fp
Quality coverage: 11.08x in Q20 bases; sum-of-contents

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that ha
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

*	1	7548:	contig	of 7548 bp	in length
*	7549	7648:	gap	of unknown length	
*	7649	31917:	contig	of 24269 bp	in length
*	31918	32017:	gap	of unknown length	
*	32018	50433:	contig	of 18416 bp	in length
*	50434	50533:	gap	of unknown length	
*	50535	95274:	contig	of 44741 bp	in length
*	95275	95374:	gap	of unknown length	
*	95375	99388:	contig	of 4614 bp	in length
*	99389	100088:	gap	of unknown length	
*	100089	109873:	contig	of 9785 bp	in length
*	109874	109973:	gap	of unknown length	
*	109974	117619:	contig	of 7646 bp	in length
*	117619	117719:	gap	of unknown length	
*	117720	128625:	contig	of 10906 bp	in length
*	128626	128725:	gap	of unknown length	
*	128726	143321:	contig	of 14796 bp	in length

Submission
 :ed (30-JUL-2003) Cell Biology, Biogen, 12 Cambridge Center,
 .ge, MA 02142, USA
 ice update by submitter
 .30, 2003 this sequence version replaced gi:2707220.
 Location/Qualifiers
 1..1239
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /cell_type="peritoneal macrophages"
 1..750
 /note="ligand in the TNF family; secreted protein"
 /codon_start=1
 /product="TWEAK"
 /protein_id="AAC53517.2"
 /db_xref="GI:33348856"
 /translation="MAARSORRRGRGPGTALLAPLVLSGLALACLGLLVVSL
 GSWATLSAEPSEQLTAEDREPELNPQTESQDVVPLEQLVRPRSAFKRKAR
 PRRAIAHREVPFGQAQAGVDGTVSGWETKINSPPRYDRQGEFTVIRAGL
 YLYICQVHDEGKAVLKLKLLVNGVLALRCLLEFSATAASSPGQLRLCQVSGLLPL
 RFGSLRIRTLFWAHLKAAFLTYFLGFQVH"
 49.8%; Score 683.4; DB 10; Length 1239;
 arity 77.0%; Pred.No. 3.8e-115;
 conservative 0; Mismatches 221; Indels 70; Gaps 9;
 :CGCCCGTCGGAGCGAGGCGGGGGCGCGGGGGGAGCGGGGCAACCGCCCTG 165
 :CGCCCGTCGGAGCGAGGCGGGGGCGCGGGGGGAGCGGGGCAACCGCCCTG 60
 :TCCCGCTCGCGTCGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 225
 :CGCCCGTCGGAGCGAGGCGGGGGCGCGGGGGGAGCGGGGCAACCGCCCTG 120
 :TCAGTTTCGGAGCGCGGCGATCGCTGCGCCAGGAGCGCTGCCAGAGGAGCTG 285
 :TCAGCGTGGGAGCTGGGCAACGCTGTGCGCCAGGAGCGCTTCTCAGAGGAGCTG 180
 :CAGAGGAGACCGAGACCGCTCGGAACCTGAATCCCGCAGACAGAAAGCCAGGAT 345
 :CAGAGGACCGCGGAGCGCCCTGAACCTGAATCCCGCAGACAGAAAGCCAGGAT 240
 :CGCCTTTCGTAACCGAATAGTTTCGGCGCTCGAGAAAGTGCACTTAAAGCGGAAA 405
 :TACCTTTCGTAACCACTAGTTCGGCGCTCGAAGAAAGTGCTCTTAAAGCGGGAAG 300
 :CGGCTCGAAGAGCGATCGAGCGCCATTTATGAAGTTTCATCCAGCCTGGACAGGAC 465
 :CGCCTCGCGAGCTATTGAGCGCCATTTATGAAGTTTCATCCAGCCTGGACAGGAT 360
 :CGAGGAGGTGTGAGCGGAGCAGTGTAGTGGCTGGGAGGAAAGCCAGAAATCAACAGC 525
 :CACAAGAGGTGTGAGCGGAGTGTAGTGGCTGGGAGGAAAGCCAGAAATCAACAGC 420
 :CGCCTTCGCGCTAACCGCGAGATCGGGAGTTTATAGTCAACCGCGCTGGGCTC 585
 :CGCCTTCGCGCTAACCGCGAGATTGGGGAATTTACAGTCAATCAGGCGTGGGCTC 480
 :ACCTGTACTGTACGTGTGACCTTTGATGAGGGAAGCGTGTCTACCTGAAGCTGGAC 645
 :ACCTGTACTGTACGTGTGACCTTTGATGAGGGAAGCGTGTCTACCTGAAGCTGGAC 540
 :TGGTGGATGGTGTGCTGGCGCTGCGCTGCGTGGAGGAATTTCTCAGCCTCGGCGC 705
 :TGGTGAACGGTGTGTGGCGCTGCGCTGCGTGGAGGAATTTCTCAGCAGCAGAGCA 600
 :CCCTCGGGCGCCAGCTCGCGCTCTGCCAGGTGTCTGGGCTGTGTGGCCCTCGCGCA 765
 :CTCCTCGGGCGCCAGCTCGGCTTGTGCCAGGTGTCTGGGCTGTGTGGCCCTCGCGCA 660
 :CTCCTCGGGATCGGACCGCTCCCTCGGCGCCATCTCAAGGCTCGCCCTTCCTC 825

Db 661 GGGTCTTCCCTTCGGATCGCACCTCCCTCCCTGGGCTCATCTTAAGGCTGCCCCCT
 QY 826 ACTACTTTCGGACTCTTCAGGTTTCACTGAGGGGCGCTTCCCTCCACAGTCGT
 Db 721 ACTACTTTCGGACTCTTTCAGGTTTCACTGAGGGGCGCTTCTCTCCAGATTCCTT
 QY 886 GCTGCCGCTCC-----CCTCGACAGCTCTCTGGGACCCCGGCTCCCTCTGCCCC
 Db 781 TTCCCTGCTCCAGGAGCATCACACACTCCCTACCCACCCCTCCCTCCAC
 QY 941 CAGCGCTTTCGCTCCAGACCTGCGCCCTCCCTCTAGAGGCTGCTGGGCGCTGT
 Db 841 C-GCTGCTCCCTTGGTCCAGTCTCTCTCC--TCAAAAGGCGAGCAGCTTGT
 QY 1001 GTTTTCCATCCACATAAATACAGTATTCCTCCACTCTTATCTTACAACTCCCCAC
 Db 898 GTTTCCTATCC-----ACAGACGATCTCTTCTTAAACATCCATCCCA
 QY 1061 ACTTCCACCTCACTAGTCTCCCAATCCCTGACCCCTTTGAGGCGCCCTGATCT
 Db 951 ACTATCCACCTCACTAGTCTCCCAAGCTTCCCAAGCCCTAC-----TTATCCC
 QY 1121 CCCCCTGGCCACAGACCCCGAGGCGATGTTCACTGACTCTCTGTGGGCAAGG
 Db 996 CCCCACACCTCACTCCCGAGCAGCTTATTAAGCTTTGTGCAC-----
 QY 1181 TCCAGAAGACCCCACTTCCAGGCTTAAGAGGGGCTTGGACCTGGCGGCGAGGAGCG
 Db 1040 -----CAGGACCTGAGTGGCTGGACCTGGTGGCAGGAGCG
 QY 1241 GACTGGGCTAGGCGAGGATTCCTCAATGTGAGGGGCGAGA-AAACAAGACAAGC
 Db 1083 ACCTGGGACTAGGCGAGGATTCCTCACTGTGAGGGGAGAGCTGGGGAACAAGC
 QY 1300 CTTGAGATTCCTGCTGATTTTAAACAGATATTTTATTTATTTATTTG
 Db 1143 CTTGGA-----TCCCTGTGGATTTTGA--AGATACTATTTTATTTATTTG
 QY 1360 ATGTTGA 1366
 Db 1197 ATGTTAA 1203
 RESULT 13
 AX180714
 LOCUS
 DEFINITION
 Sequence 1 from Patent WO0145730.
 ACCESSION
 AX180714
 VERSION
 AX180714.1 GI:15132570
 KEYWORDS
 synthetic construct
 ORGANISM
 synthetic construct
 artificial sequences.
 REFERENCE
 1
 AUTHORS
 Wiley, S.R.
 TITLE
 Tweak receptor
 JOURNAL
 Patent: WO 0145730-A 1 28-JUN-2001;
 IMMUNEX CORPORATION (US)
 FEATURES
 Location/Qualifiers
 source
 1..898
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 52..873
 /note="unnamed protein product; human TWEAK fusic
 construct"
 /codon_start=1
 /translation="MATGSRSTSLLLAFGLLCLPWLQEGSATSSDRMKQI

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.
 .c search, using sw model
 File 7, 2004, 17:30:19 ; Search time 524.669 Seconds
 (without alignments)
 11117.066 Million cell updates/sec

-09-245-198A-3
 73
 atgcattgttagacttga.....gacaaatgtgataaatgg 1373

ENTITY NUC

Gap 10.0, Gapext 1.0

73863 seqs, 2124099041 residues

is satisfying chosen parameters: 6747726

hth: 0

hth: 2000000000

Minimum Match 0%
 Maximum Match 100%
 string first 45 summaries

Geneseq 29Jan04: *
 geneseqn1980s: *
 geneseqn1990s: *
 geneseqn2000s: *
 geneseqn2001as: *
 geneseqn2001bs: *
 geneseqn2002s: *
 geneseqn2003as: *
 geneseqn2003bs: *
 geneseqn2003cs: *
 geneseqn2004s: *

the number of results predicted by chance to have a
 value greater than or equal to the score of the result being printed,
 as determined by analysis of the total score distribution.

SUMMARIES

Seq	Length	DB	ID	Description
1.0	1373	2	AAV18600	AAV18600 Homo sapi
1.0	1364	6	ABK34881	ABK34881 Human CDN
1.5	1421	2	AAK56000	AAK56000 Human tum
1.2	1353	3	AAA49717	AAA49717 Human PRO
1.2	1353	6	ABK40255	ABK40255 cDNA enco
1.6	1306	7	ACC57587	ACC57587 Polynucle
1.6	1306	7	ACC57901	ACC57901 Human TWE
1.6	1306	9	ADC35205	ADC35205 Human CDN
1.3	1236	2	AAV47613	AAV47613 TNF relat
1.3	1236	4	AAV4350	AAV4350 Human TRE
1.7	1030	2	AAK23424	AAK23424 Human TNR
1.8	1239	9	ADC97713	ADC97713 Murine FL
1.8	898	4	AAK03964	AAK03964 Expressio
1.8	1168	2	AAV18599	AAV18599 Mus muscu
1.8	701	2	AAK23425	AAK23425 Mouse TNR
1.4	493	8	ACH34013	ACH34013 Human end
1.5	408	7	ABX37032	ABX37032 Bovine ES
1.9	282	2	AAT22190	AAT22190 Human gen
1.1	195	6	ABK29540	ABK29540 Colon ade
1.1	412	9	ADB56326	ADB56326 Toxicity
1.7	264	7	ABX52254	ABX52254 Bovine ES
1.1	114955	2	AAK53491	AAK53491 Human ade
1.9	3163	9	ADC87060	ADC87060 Human GPC

Seq	Length	DB	ID	Description
24	65.4	4.8	1064	6 AAT09678 Hu
25	63.8	4.6	5452	9 ADC86736 Hu
26	62.2	4.5	12733	6 ABK98631 Ve
27	62.2	4.5	12733	8 ACD13882 L.
28	62.2	4.5	12739	8 ABK98592 Ve
29	62.2	4.5	12739	8 ACD13843 Fl
30	61.4	4.5	1117	9 ADC86688 Hu
31	61	4.4	1337	2 AAZ17263 Hu
32	60.6	4.4	1000	3 AAA02484 Hu
33	60	4.4	60	6 ABN58848 Hu
34	60	4.4	60	6 ABN41049 Hu
35	60	4.4	60	6 ABN58593 Hu
36	60	4.4	60	6 ABN58591 Hu
37	60	4.4	60	6 ABN58849 Hu
38	59.6	4.3	3133	9 ADC86738 Hu
39	59.2	4.3	10732	3 AAA10594 Ge
40	58	4.2	1218	3 AAA02488 Hu
41	56	4.1	1065	6 AAT09682 Hu
42	54.8	4.0	1017	7 AAD36876 S.
43	54.8	4.0	2000	7 ADA71938 Ri
44	54.8	4.0	29870	7 AAD36874 St
45	54.4	4.0	600	6 ABQ52497 Ol

ALIGNMENTS

RESULT 1
 ID AAV18600 standard; cDNA; 1373 BP.
 AC AAV18600;
 DT 21-JUL-1998 (first entry)
 XX Homo sapiens tumour necrosis factor related ligand (TRELL) gene.
 DE TRELL; tumour necrosis factor related ligand; tnf; treatment; can
 KW autoimmune disease; immune system; stimulation; suppression;
 KW graft rejection; ds.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 1..852
 FT /tag= a
 FT /note= "tumour necrosis factor related ligand"
 XX W09805783-A1.
 PD 12-FEB-1998.
 PF 07-AUG-1997; 97WO-US013945.
 PR 07-AUG-1996; 96US-0023541P.
 PR 18-OCT-1996; 96US-0028515P.
 PR 18-MAR-1997; 97US-0040820P.
 XX (BIOJ) BIOGEN INC.
 PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.
 XX Chicheportiche Y, Browning JL;
 WPI: 1998-145619/13.
 P-PSDB; AAW47525.
 XX Tumour necrosis factor related ligand - useful for, e.g. treating
 PT auto-immune disease and immune responses to tissue grafts.
 XX Claim 2; Page 48-50; 69pp; English.
 XX The sequence is that encoding human tumour necrosis factor relate
 CC (TRELL). TRELL or active fragments can be included with a carrier

XX
PA (GEMV) GENETICS TNST TNC

Location/Qualifiers

58..807
/*tag= a
58..177
/*tag= b
178..804
/*tag= c

2.

99WO-US028565.

98US-0113296P.
99WO-US005028.
99US-0130232P.
99US-0131445P.
99US-0134287P.
99US-0144758P.
99US-0145698P.
99WO-US021090.
99WO-US021547.

TECH INC.

Goddard A, Godowski PJ, Gurney AL, Marsters SA;
tbi RM, Wood WI;

58/38.
138.

ion to inhibit neoplastic cell growth or for treating tumor
risers polypeptides PRO179, PRO207, PRO320, PRO219, PRO221,
3, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866.

3; 172pp; English.

quence is that of cDNA clone DNA30879-1152 (ATCC 209358)
PRO207 (see AAY95338), which shows homology to several
tumour necrosis factor family, especially human
3.4). The cDNA was identified in a foetal kidney cDNA
man identification of an expressed sequence tag with
man Apo-2 ligand. A claimed method for inhibiting the
tumour cell comprises exposing the tumor cell to PRO179,
PRO219, PRO221, PRO224, PRO328, PRO301, PRO526, PRO362,
or PRO866 (see AAY95337-49), their agonists or chimeric
incorporating them. The tumour is especially a cancer
breast, ovarian, renal, colorectal, uterine, prostate,
and central nervous system cancer, melanoma and leukaemia.
encoding PRO179 etc. are used in the recombinant production
our polypeptides

BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

96.2%; Score 1320.2; DB 3; Length 1353;
arity 99.8%; Pred. No. 3.3e-271;
nservative 0; Mismatches 3; Indels 0; Gaps 0;

1CTCGGCTCCCGGATGGGGGCGGTAGGAGCAGACAGCCCCCGCCCCCATG 108
1CCTCGGCTCCCGGATGGGGGCGGTAGGAGCAGACAGCCCCCGCCCCCATG 60
1CCGTCCGAGCCAGAGCGGAGGGGCGCGCGGGGAGCGGGCACCGCCCTGCTG 168
1CCGTCCGAGCCAGAGCGGAGGGGCGCGCGGGGAGCGGGCACCGCCCTGCTG 120
1CTCGGCTCCCGGATGGGGGCGGTAGGAGCAGACAGCCCCCGCCCCCATG 228
1GCTCGGCTCCCGGATGGGGGCGGTAGGAGCAGACAGCCCCCGCCCCCATG 180
1TTTGGGAGCCGGGATCGCTGCGCCAGGAGCCTGCCAGGAGGAGCTGGTG 288

181 GTCAGTTTGGGGAGCCGGGCATCGCTGTCGCGCCAGAGAGCTGCCAGGAGAGCT
289 GCAGAGGAGGAGCCAGGACCGTCGGAACCTGAATCCCCAGACAGAAAGCCAGGA
241 GCAGAGGAGGAGCCAGGACCGTCGGAACCTGAATCCCCAGACAGAAAGCCAGGA
349 GCGCCTTTTCTGAAACCGACTAGTTTCGGCCTCGCAGAAAGTGCACCTAAAGCCGGAA
301 GCGCCTTTTCTGAAACCGACTAGTTTCGGCCTCGCAGAAAGTGCACCTAAAGCCGGAA
409 CGGGCTCGAAGAGCGATCGCAGCCCAATTATGAAGTTTCATCCAGACCTGACAGGA
361 CGGGCTCGAAGAGCGATCGCAGCCCAATTATGAAGTTTCATCCAGACCTGACAGGA
469 GCGCAGGAGGTGTGACAGGGACAGTGAGTGGCTGGGAGGAAGCCAGAAATCAACAG
421 GCGCAGGAGGTGTGACAGGGACAGTGAGTGGCTGGGAGGAAGCCAGAAATCAACAG
529 AGCCCTCTGCGCTACAAACCGCCAGATCGGGGAGTTTTATAGTACCCGGGCTGGGCT
481 AGCCCTCTGCGCTACAAACCGCCAGATCGGGGAGTTTTATAGTACCCGGGCTGGGCT
589 TACCTGTACTGTGAGTGCATTTTGTATGAGGGAAGGCTGTCTACCTGAAGCTGGA
541 TACCTGTACTGTGAGTGCATTTTGTATGAGGGAAGGCTGTCTACCTGAAGCTGGA
649 CTGGTGGATGGTGTGCTGGCCCTGCGCTGCTGAGGAAATTTCTCAGCCACTGCGGC
601 CTGGTGGATGGTGTGCTGGCCCTGCGCTGCTGAGGAAATTTCTCAGCCACTGCGGC
709 TCCTTCGGGCCCCAGCTCCGCTCTCCAGGTTCTGGGCTGTGGGCTGTGGCCCTGCGGC
661 TCCTTCGGGCCCCAGCTCCGCTCTCCAGGTTCTGGGCTGTGGGCTGTGGCCCTGCGGC
769 TCCTTCGGGCCCCAGCTCCGCTCTCCAGGTTCTGGGCTGTGGGCTGTGGCCCTGCGGC
721 TCCTTCGGGCCCCAGCTCCGCTCTCCAGGTTCTGGGCTGTGGGCTGTGGCCCTGCGGC
829 TACTTCGGACTCTTCCAGGTTCTCAGGAGGCGCTGCTCCCCACAGCTCGTCCCA
781 TACTTCGGACTCTTCCAGGTTCTCAGGAGGCGCTGCTCCCCACAGCTCGTCCCA
889 GCGGCTCCCTCGACAGCTCTCTGGGACCCCGGCTCCCTCTGCCCCACCCCTCAGC
841 GCGGCTCCCTCGACAGCTCTCTGGGACCCCGGCTCCCTCTGCCCCACCCCTCAGC
949 CTTTGTCTCAGACCTCCCTCCCTCTAGAGGCTGCTGGGCTGTTCAGTGTCTT
901 CTTTGTCTCAGACCTCCCTCCCTCTAGAGGCTGCTGGGCTGTTCAGTGTCTT
1009 TCCACATAAATACAGTATTTCCACCTCTTATCTTACAACTCCCCACCGCCACTC
961 TCCACATAAATACAGTATTTCCACCTCTTATCTTACAACTCCCCACCGCCACTC
1069 CTTCACTAGTCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGATCTGACTCCCCC
1021 CTTCACTAGTCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGATCTGACTCCCCC
1129 GCCACAGCCCCCAGGCAATGTGTTCATCTACTCTGTGGGCAAGATGGGTCCA
1081 GCCACAGCCCCCAGGCAATGTGTTCATCTACTCTGTGGGCAAGATGGGTCCA
1189 ACCCACTTCAGGCACTAAGAGGGGCTGACCTGGGCGCAGGAAGCAAGAGACTT
1141 ACCCACTTCAGGCACTAAGAGGGGCTGACCTGGGCGCAGGAAGCAAGAGACTT
1249 CTAGGCCAGAGTTCCTCAATGTGAGGGGCGAGAAACAAGCAAGCTCTCTCCCTTG
1201 CTAGGCCAGAGTTCCTCAATGTGAGGGGCGAGAAACAAGCAAGCTCTCTCCCTTG
1309 TTCCCTCTGAGTTCCTCAATGTGAGGGGCGAGAAACAAGCAAGCTCTCTCCCTTG
1261 TTCCCTCTGAGTTCCTCAATGTGAGGGGCGAGAAACAAGCAAGCTCTCTCCCTTG

GG 1373
||
GG 1325

ndard; cDNA; 1353 BP.

(first entry)

g human PRO207 polypeptide.

benign tumour; malignant tumour; lymphoid malignancy;
neural disorder; stromal disorder; blastocoele disorder;
disorder; immune disorder; angiogenic disorder;
; cytostatic; neuroprotective; gene; ss.

11.

2000WO-US003565.

99WO-US005028.
99US-0123972P.
99US-0133459P.
99WO-US012252.
99US-0140850P.
99US-0140653P.
99US-0144758P.
99US-0145698P.
99US-0146222P.
99US-0149395P.
99US-0151889P.
99WO-US020111.
99WO-US021090.
99WO-US028313.
99WO-US028301.
99WO-US028634.
2000WO-US000219.

ITECH INC.

Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
Wood WI;

567/26.
129.

nucleic acids encoding PRO polypeptides, useful for treating
ignant tumors, leukemias and lymphoid malignancies,
angiogenic and immunologic disorders.

f 3; 302pp; English.

vention relates to the isolation of novel human PRO
and the polynucleotide sequences encoding them. The PRO
agonists, antagonists or anti-PRO antibodies are useful for
gn or malignant tumours (e.g. renal, kidney, bladder,
leukaemias and lymphoid malignancies, other disorders such
glial, astrocytal, hypothalamic, glandular, macrophagal,
blastocoele disorders, inflammatory, immune and angiogenic
ie polynucleotide sequences are also useful in gene therapy.
0288 encode for the human PRO polypeptides of the invention

BP; 257 A; 443 C; 389 G; 264 T; 0 U; 0 Other;

Query Match		96.2%;	Score 1320.2;	DB 6;	Length 1353;
Best Local Similarity		99.8%;	Pred. No. 3	3e-271;	
Matches 1322;		Conservative	0;	Mismatches	3; Indels 0;
QY	49	CGATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCC			
DB	1	CGATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCC			
QY	109	GCGGCCCTCGGAGCCAGAGGGCGGAGGGGGCGCCGGGGGAGCCGGGACCGCCC			
DB	61	GCGGCCCTCGGAGCCAGAGGGCGGAGGGGGCGCCGGGGGAGCCGGGACCGCCC			
QY	169	GTCCGCTCGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG			
DB	121	GTCCGCTCGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG			
QY	229	GTGAGTTTGGGAGCGGGGATCGTGTGCGGCCAGGAGCCTGCGCCAGAGAGC			
DB	181	GTGAGTTTGGGAGCGGGGATCGTGTGCGGCCAGGAGCCTGCGCCAGAGAGC			
QY	289	GCAGAGGAGGACAGGACCCGTCGGAATCGAATCCCGACAGAGAGAAAGCCAGG			
DB	241	GCAGAGGAGGACAGGACCCGTCGGAATCGAATCCCGACAGAGAGAAAGCCAGG			
QY	349	GCGCCTTTCCTGAACCGACTAGTTCGGCTCGCAGAAGTGCACCTAAAGCCGGA			
DB	301	GCGCCTTTCCTGAACCGACTAGTTCGGCTCGCAGAAGTGCACCTAAAGCCGGA			
QY	409	CGGGCTCGAAGAGCGATCGCAGCCCTATTAAGTTTCATCCAGACCTGGACAGG			
DB	361	CGGGCTCGAAGAGCGATCGCAGCCCTATTAAGTTTCATCCAGACCTGGACAGG			
QY	469	GCGCAGGAGGTTGCGACGGGACAGTGTGCTGGGAGGAGGAGGAGGAGGAGG			
DB	421	GCGCAGGAGGTTGCGACGGGACAGTGTGCTGGGAGGAGGAGGAGGAGGAGG			
QY	529	AGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGG			
DB	481	AGCCCTCTGCGCTACAACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGG			
QY	589	TACCTGTACTGTGAGTGTGCTGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGG			
DB	541	TACCTGTACTGTGAGTGTGCTGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGG			
QY	649	CTGTGTGATGTGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG			
DB	601	CTGTGTGATGTGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG			
QY	709	TCCCTCGGGGCGCCAGCTCGGCTCTGCGAGGTTCTGGGCTGTGGGCTGTGGGCT			
DB	661	TCCCTCGGGGCGCCAGCTCGGCTCTGCGAGGTTCTGGGCTGTGGGCTGTGGGCT			
QY	769	TCCTCCCTCGGGATCCGACCCCTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTC			
DB	721	TCCTCCCTCGGGATCCGACCCCTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTC			
QY	829	TACTTCGAGCTCTTCCAGGTTCACTGAGGGGCGCTGGTCTCCCGACAGTCTCCG			
DB	781	TACTTCGAGCTCTTCCAGGTTCACTGAGGGGCGCTGGTCTCCCGACAGTCTCCG			
QY	889	GCGGCTCCCTCGACAGCTCTCTGGGCA CCGGTTCCCTCTGGCCCAACCTCAGC			
DB	841	GCGGCTCCCTCGACAGCTCTCTGGGCA CCGGTTCCCTCTGGCCCAACCTCAGC			
QY	949	CTTTGTCCAGACCTGCGCTCCCTCTAGAGGCTGCTGGGCTGCTGGGCTGCTGG			
DB	901	CTTTGTCCAGACCTGCGCTCCCTCTAGAGGCTGCTGGGCTGCTGGGCTGCTGG			
QY	1009	TCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCGCACCGCCACTC			
DB	961	TCCACATAAATACAGTATTCCCACTCTTATCTTACAACTCCCGCACCGCCACTC			

CTAGCTCCCCNAATCCCTGACCCCTTGAGGGCCCCAGTGATCTCGACTCCCCCTCG 1128
 CTAGCTCCCCNAATCCCTGACCCCTTGAGGGCCCCAGTGATCTCGACTCCCCCTCG 1080
 AGACCCCCACAGGCGATTGTGTTCACTGTACTCTGTGGGCAAGATGGGTCCAGAAG 1188
 AGACCCCCACAGGCGATTGTGTTCACTGTACTCTGTGGGCAAGATGGGTCCAGAAG 1140
 ACTTTCAGGCACCTAAGAGGGGCTGGACCTGGCGCGCAGGAAGCCAAAGAGACTGGGC 1248
 ACTTTCAGGCACCTAAGAGGGGCTGGACCTGGCGCGCAGGAAGCCAAAGAGACTGGGC 1200
 CCAGGAGTCTCCCAATGTGAGGGCGAGAAACAAGACAAGCTCTCTCCCTTCAGAA 1308
 CCAGGAGTCTCCCAATGTGAGGGCGAGAAACAAGACAAGCTCTCTCCCTTCAGAA 1260
 CTGTGGATTTTAAAAACAGATATTATTTTATTATTATTTGTCACAAAATGTTGATA 1368
 CTGTGGATTTTAAAAACAGATATTATTTTATTATTATTTGTCACAAAATGTTGATA 1320

lard; DNA: 1306 BP.

(first entry)

encoding tumour necrosis factor superfamily member.

tumour necrosis factor; osteopathic; bone; gene; ds.

12.

002WO-US033022.

:001US-0329393P.

3-JEWISH HOSPITAL.

'F, Teitelbaum SL;

46/40.

comprising a core, and at least one external loop, useful processes of bone formation or inhibiting bone resorption, treatments for disease or condition characterized by loss

ige 66-67; 78pp; English.

sequence is that of a polynucleotide encoding a non-RANKL tumour necrosis factor (TNF) superfamily. The invention of naturally-occurring proteins that contain one or more of the ice loops of RANKL (see ABR42066-70) in combination with a protein core obtained from a non-RANKL member of the TNF family also provided are polynucleotides encoding such proteins. In addition, the invention provides polynucleotides that encode a mimic of RANKL, acting as mimics of RANKL. They can be used to modulate bone formation by either inhibiting bone resorption or inducing bone formation, thus providing treatment for diseases or conditions such as loss of bone mass.

BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;

Query Match	93.6%	Score 1285	DB 7	Length 1306
Best Local Similarity	100.0%	Pred. No. 1e-263		
Matches 1285	Conservative 0	Mismatches 0	Indels 0	G
QY	89	CACAGCCCCCGCCCCCATGTCGCGCCCGCTCGAGCCAGAGCGAGGGGCGCGCG		
DB	1	CACAGCCCCCGCCCCCATGCGCGCCCGCTCGAGCCAGAGGGGCGCGCG		
QY	149	AGCCGGGCAACGGCCCTGCTGCTCCGCTCGCGCTGGGCTTGGGCTCTGGCGCTCGCGCT		
DB	61	AGCCGGGCAACGGCCCTGCTGCTCCGCTCGCGCTGGGCTCTGGGCTCTGGCGCTCGCGCT		
QY	209	TCGGCTCTCTGCTGGCGCGTGTGAGTTTGGGAGCCGGGCATCGCTGTCCGCCAG		
DB	121	TCGGCTCTCTGCTGGCGCGTGTGAGTTTGGGAGCCGGGCATCGCTGTCCGCCAG		
QY	269	CTGCCACAGGAGAGCTGCTGGCAGAGGAGACACAGGACCCGTCGGAACTCGAATCCC		
DB	181	CTGCCACAGGAGAGCTGCTGGCAGAGGAGACACAGGACCCGTCGGAACTCGAATCCC		
QY	329	CAGAAGAAAGCCAGGATCTCTGGCGCTTCTTCTGAAACCGACTAGTTTCGGCCTCGCAGA		
DB	241	CAGAAGAAAGCCAGGATCTCTGGCGCTTCTTCTGAAACCGACTAGTTTCGGCCTCGCAGA		
QY	389	CACCTAAAGGCGGAAACACCGGCTCGAAGAGCGATCGCAGGCCCATTAATGAAGTT		
DB	301	CACCTAAAGGCGGAAACACCGGCTCGAAGAGCGATCGCAGGCCCATTAATGAAGTT		
QY	449	CACGACCTGACACAGGACGAGCGCAGGACGAGTGTGGACGGGACAGTGAAGTGGCTGG		
DB	361	CACGACCTGACACAGGACGAGCGCAGGACGAGTGTGGACGGGACAGTGAAGTGGCTGG		
QY	509	AAGCCAGAATCAACAGCTCCAGCGCTCTGCGCTACACCGCCAGATCGGGGAGTTT		
DB	421	AAGCCAGATCAACAGCTCCAGCGCTCTGCGCTACACCGCCAGATCGGGGAGTTT		
QY	569	TCACCCGGGCTGGGCTCTACTACTCTGCTGAGGTGCATTTTGTATGAGGGGAAG		
DB	481	TCACCCGGGCTGGGCTCTACTACTCTGCTGAGGTGCATTTTGTATGAGGGGAAG		
QY	629	TCTACTGTAAGCTGGAATTTGCTGGTGGATGTTGTGTGTGGCCCTGCGCTGCTGGAG		
DB	541	TCTACTGTAAGCTGGAATTTGCTGGTGGATGTTGTGTGTGGCCCTGCGCTGCTGGAG		
QY	689	TCTCAGCCACTGCGGCGAGTTCCTCTCGGCGCCACAGCTCCGCTCTGCGCAGGCTCTCT		
DB	601	TCTCAGCCACTGCGGCGAGTTCCTCTCGGCGCCACAGCTCCGCTCTGCGCAGGCTCTCT		
QY	749	TGTTGGCCCTGCGGCGAGGCTCTCCCTCGGGAATCCGACCCCTCCCTGCGGCCCAT		
DB	661	TGTTGGCCCTGCGGCGAGGCTCTCCCTCGGGAATCCGACCCCTCCCTGCGGCCCAT		
QY	809	AGGCTGCCCCCTTCTCACTACTTCGGAATCTTCCAGGTTTCACTGAGGGGCGCCTG		
DB	721	AGGCTGCCCCCTTCTCACTACTTCGGAATCTTCCAGGTTTCACTGAGGGGCGCCTG		
QY	869	CCCCACAGTCTGCCAGGCTCGCGCTCCCTCCACAGAGCTCTCTGCGGACACCGGCTC		
DB	781	CCCCACAGTCTGCCAGGCTCGCGCTCCCTCCACAGAGCTCTCTGCGGACACCGGCTC		
QY	929	CTGCCCAACCTCAGCGCTCTTTGCTCCAGACCTGCGCCCTCCCTCTAGAGGCTGC		
DB	841	CTGCCCAACCTCAGCGCTCTTTGCTCCAGACCTGCGCCCTCCCTCTAGAGGCTGC		
QY	989	GCCTGTTACGTTGTTTTCATCCACATTAATAAGTAATTCACATCTTATCTTAC		
DB	901	GCCTGTTACGTTGTTTTCATCCACATTAATAAGTAATTCACATCTTATCTTAC		
QY	1049	CCCCCACCGCCACTCTCCACTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCC		
DB	961	CCCCCACCGCCACTCTCCACTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCC		
QY	1109	GTGATCTCGACTCCCCCTTGGGCCACACAGACCCCGCAGGGCATTTGTTCACTGTACTCT		

19

|||||ATCTCGAGCTCCCTGGCCACAGACCCCGAGGCAATTGTTCACTGTACTCTGTG 1080
AAGATGGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGCA 1228
AAGATGGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGCA 1140
AGCCAAAGAGACTGGGCTAGGCAGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1288
AGCCAAAGAGACTGGGCTAGGCAGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1200
AGCTCCCTCCCTGAGAAATCCCTGTGATTTTAAACAGATATATTTTATTTATTT 1348
AGCTCCCTCCCTGAGAAATCCCTGTGATTTTAAACAGATATATTTTATTTATTT 1260
GTGACAAAATGTTGATAAATGG 1373
GTGACAAAATGTTGATAAATGG 1285

standard; cDNA; 1306 BP.

(first entry)

coding sequence.

; tumour necrosis factor; ligand; cytostatic;
factor; osteopathic; gene; ss.

Location/Qualifiers

18..767

/tag= a
/product= "Human TWEAK"

-A2.

2002WO-US023782.

2001US-0307838P.

GENOME SCI INC.

Rosen CA;

659/40.

2315.

oligomeric complex having a first polypeptide member of the
tumor necrosis factor (TNF) ligand family, and a second different member
of the TNF ligand family, useful for treating cancer, osteoporosis or an
autoimmune disease.

Page 367-368; 388pp; English.

sequence is that of a polynucleotide encoding human TWEAK.
The invention relates to compositions comprising heterotrimeric complexes
of tumor necrosis factor (TNF) ligand family members, and their use in
the prevention and treatment of disease. In one embodiment,
the heterotrimeric complex comprises full-length or extracellular
TWEAK and full-length or extracellular portions of other TNF
family members, preferably VEGF or VEGF-SV. The heterotrimeric
complex is useful for treating an autoimmune disease,
osteoporosis, and particularly for inhibiting cancer cell
proliferation, or inducing apoptosis of

XX
SQ Sequence 1306 BP; 247 A; 434 C; 368 G; 257 T; 0 U; 0 Other;
Query Match 93.6%; Score 1285; DB 7; Length 1306;
Best Local Similarity 100.0%; Pred. No. 1e-263;
Matches 1285; Conservative 0; Mismatches 0; Indels 0;
QY 89 CACAGCCCCCGCCCCCATGGCCGCTCGAGCCGAGAGGCGGAGGGGCGCGCG
Db 1 CACAGCCCCCGCCCCCATGGCCGCTCGAGCCGAGAGGCGGAGGGGCGCGCG
QY 149 AGCGGGGCAACCGCCCTGCTGGTTCGCGCTCGCGCTGGGCTGGGCTGGGCTGGC
Db 61 AGCGGGGCAACCGCCCTGCTGGTTCGCGCTCGCGCTGGGCTGGGCTGGGCTGGC
QY 209 TGGGCTCTCTGCTGGCCGCTGGTTCAGTTTGGGGAGCGGGCATCGCTGTCGGCCCA
Db 121 TGGGCTCTCTGCTGGCCGCTGGTTCAGTTTGGGGAGCGGGCATCGCTGTCGGCCCA
QY 269 CTGCCAGGAGGAGCTGGTGGCAGAGGAGGAGCCAGGACCCGTCGGAATGAATCC
Db 181 CTGCCAGGAGGAGCTGGTGGCAGAGGAGGAGCCAGGACCCGTCGGAATGAATCC
QY 329 CAGAGAAAGCCAGGATCTGCGCTTCTGCTGAAACCGACTAGTTGGGCTGGGAG
Db 241 CAGAGAAAGCCAGGATCTGCGCTTCTGCTGAAACCGACTAGTTGGGCTGGGAG
QY 389 CACCTAAAGGCGGAAACACAGGCTCGAAGAGCGATCGAGCCCATTAATGAGT
Db 301 CACCTAAAGGCGGAAACACAGGCTCGAAGAGCGATCGAGCCCATTAATGAGT
QY 449 CAGACCTGGACAGGACGCGAGCGAGGAGGTGTGGACGGGACAGTGAAGTGGCTG
Db 361 CAGACCTGGACAGGACGCGAGCGAGGAGGTGTGGACGGGACAGTGAAGTGGCTG
QY 509 AAGCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGAGTT
Db 421 AAGCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAACCGCCAGATCGGGAGTT
QY 569 TCACCCGGGCTGGGCTCTACTACTGTACTGTGTCAGGTGCACCTTTGATGAGGGAA
Db 481 TCACCCGGGCTGGGCTCTACTACTGTACTGTGTCAGGTGCACCTTTGATGAGGGAA
QY 629 TCTACCTGAAGCTGGAATGCTGTGGTGGTGTGTGTCGGCCCTGGGCTGGCTGGAA
Db 541 TCTACCTGAAGCTGGAATGCTGTGGTGGTGTGTGTCGGCCCTGGGCTGGCTGGAA
QY 689 TCTCAGCCACTGGCGGCGAGTTCCCTGGGGCCCCAGCTCGGCTCTGCCAGGTGTC
Db 601 TCTCAGCCACTGGCGGCGAGTTCCCTGGGGCCCCAGCTCGGCTCTGCCAGGTGTC
QY 749 TGTGGGCTCTGGCGGCGAGGTCTCTCCCTGGGATCGGACCCCTCCCTGGGGCCCA
Db 661 TGTGGGCTCTGGCGGCGAGGTCTCTCCCTGGGATCGGACCCCTCCCTGGGGCCCA
QY 809 AGGCTGCCCCCTCTCTCCTACCTACTTCCAGCTTCTCCAGTTTCACTGAGGGGCTCT
Db 721 AGGCTGCCCCCTCTCTCCTACCTACTTCCAGCTTCTCCAGTTTCACTGAGGGGCTCT
QY 869 CCCCACAGTGTCTCCAGGCTGGGCTGGCTGGCTGGACAGCTCTCTGGGACCCCGGT
Db 781 CCCCACAGTGTCTCCAGGCTGGGCTGGCTGGCTGGACAGCTCTCTGGGACCCCGGT
QY 929 GTGCCCCCACCCTCAGCCGCTCTTTGCTCCAGACCTTCCCTCTCTAGAGGCTCTG
Db 841 GTGCCCCCACCCTCAGCCGCTCTTTGCTCCAGACCTTCCCTCTCTAGAGGCTCTG
QY 989 GCCTGTTACAGTGTGTTTCCATCCCAATAAATACAGTATTCCTCTTATCTTA
Db 901 GCCTGTTACAGTGTGTTTCCATCCCAATAAATACAGTATTCCTCTTATCTTA
QY 1049 CCCCACAGGCGGCTCTCTCCTACCTCTCAGTGTCCCAATCCCTGAGCCCTTGAAGCT

CCACCGCCCACTCTCCGACCTCACTAGCTCCCCAAATCCCTGACCCCTTTGAGGCCCCCA 1108
 CCACCGCCCACTCTCCACCTCACTAGCTCCCCAAATCCCTGACCCCTTTGAGGCCCCCA 1020
 ATCTCGACTCCCCCTGGCCACAGACACCCAGGGCATTGTGTTCACTGTACTCTGTG 1168
 ATCTCGACTCCCCCTGGCCACAGACACCCAGGGCATTGTGTTCACTGTACTCTGTG 1080
 AAGGATGGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGCTGACCTGGCGGCA 1228
 AAGATGGGTCCAGAGACCCCACTTCAGGCACCTAAGAGGGCTGACCTGGCGGCA 1140
 AGCCAAAGACACTGGGCCTAGGCAGAGATTCCAAATGTGAGGGCGAGAAACAAG 1288
 AGCCAAAGACACTGGGCCTAGGCAGAGATTCCAAATGTGAGGGCGAGAAACAAG 1200
 AGTCTCTCCCTTCAGAAATCCCTGTGGATTTTAAACAGATATATTTTATTATT 1348
 AGTCTCTCCCTTCAGAAATCCCTGTGGATTTTAAACAGATATATTTTATTATT 1260
 GTGACAAAATGTTGATAAATGG 1373
 GTGACAAAATGTTGATAAATGG 1285

ndard; cDNA; 1236 BP.

(first entry)

endothelium proliferative agent gene.

epithelium proliferative agent; TREPA; wound healing; cancer; ing; vascularisation; apoptosis; autoimmune; birth control.

```
Location/Qualifiers
1. .750
/*tag= a
/product= "TREPA"
```

98WO-US002859.

97US-00798692.
98US-00021706.

IT LAB.

'255/38.
'745.

oleic acid encoding TREPA - useful for diagnosis and autoimmune disease, tumours and inflammation.

je 123-4; 142pp; English.

ed endothelium proliferative agent (TRPRA), or its agonists, are used to treat a deficit of TRPRA, e.g. to healing or tissue grafting, by promoting vascularisation. The apoptosis for treating cancer and eliminating autoreactive as adjunct to cancer chemotherapy or antiviral treatment. s can also be used to target cytotoxic agents or for

dard; cDNA: 1236 bp.

TNF related endothelium proliferative agent) cDNA.

Location/Qualifiers

1. .750

```
/*tag= a
/product= "Human TREPA (TNF related endothelium
proliferative agent)"
```

97US-00798692.

98US-00021706.

'60/29.

191.

genesis in mammal at desired sites for promoting wound
ministering soluble fragment of extracellular domain of
factor related endothelium proliferative agent protein.

73-74; 53pp; English.

vention relates to extracellular signal molecules,

Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 U; 0 Other;

Query Match 89.3%; Score 1236.4; DB 4; Length 1236;
Best Local Similarity 99.5%; Pred. No. 2.9e-251;
Matches 1230; Conservative 0; Mismatches 6; Indels 0.

QY 106 ATGGCGGCCGTGGAGCCAGAGCGGAGGGGGCGCCGGGGGAGCCGGGCA CCG(

Db
1 ATGGCCGCCCGTCCGAGCCAGAAGCGGAGGGGGCGCGGGGGAGACCGGGCACCGG

QY 166 CTGGTCCGCTCGGCTGGGCCTGGGCCTGGGCCTGCCCTCGGCCCTCCTGC

[illegible]

0v 226 ATTACHED TO THE RECORDS OF THE

121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632

DD 121 GTGGTCAGTTTGGGGAGCCGGGGCATCGCTGTCCGCCCCAGGAGCCTGCCCCAGGAGG

286 GTGGCAGGAGGACCGAGACCGGTGGAACTGAATCCCCAGACAGAAGAAAGCCF QY

Db 181 GTGGCAGAGGAGGACCAGGACCCGTCGGAACTGAATCCCCAGACAGAGAAGCCCF

346 CCTGGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAAGTGCACTAAAGGCCG

Dbb 241 CCTGGCCCTTCCCTGAACCGACTAGTTCGGCCCTCGAAGAAGTGCCACCTAAGGCCG

QY 406 ACACGGGCTCGAAGAGCGGATCGCAGGCCATTATGAGTTTCATCCACGACCTGCGAC

b6
301 ACACGGCTCGAGACCCATGCGCCCCCAATTATTCAGAACCTAATTTTAACTCCTTGACTG
|||||

[illegible]

466 GGAGCGCAGGCAGGTGTGGACGGGACAGTGGTGGCTGGGAGGAGCCAGATCA

Db
361 GGAGCGCAGGCAGGTGTGGACGGGACAGTGAGTGGCTGGGAGGAAGCCAGATCAA

QY 526 TCCAGCCCTCTGGCTACAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGG

db
421 TCCAGCCCTCTGGCTACACGCCAGATCGGGAGTTTATAGTCACCCGGGCTGG

586 TACTACCTGTACTGTCAAGTGCACTTTGATGACGGGAAGCGTGTCTACCTGAAGCT

db 481 TACTACCTGTACTGTCAAGTGCACTTTGATGAGGGGAAGCGTGTCTACCTGAAGCT

646 TTGCTGGTGGATGGTGCTGGCGCCCTGCCCTGCCTGGCGCAATTTCACAGCAGCACAC

5 41

341 TTGCTGGAGGAATTCAGCCACTGC

706 AGTTCCTTCGGGCCCCAGCTCCGGCTCTGCCAGGTGTCTGGGCTGTGGCCCCTCGG

Db
601 AGTTCCTCGGGCCCGAGTCCGCCCTCTGCCAGGTGTCTGGGCTGTTGGCCCCCTGGC

766 GGGTCCTCCCTGCGGATCGCACCTCCCTGGGCCCATCTCAAGGCTGCCCCCTT(

661 GGGTCCTCCCTGGGATCGGCACCCCTCCCCTGGGGCCCATCTCAAGGCTGCCCCCTT(

826 ACCTACTTTCGGACTCTTCCAGGTTCACGTAGGGGCCCTGGTCTCCCCACACAGTCGCTC

[illegible][illegible]

886 GCAGCGGCTCCCTCGACAGCTCTCTGGGCACCCGGTCCCTCTGCCCCACCCCTG

db 781 GCTGCCGGCTCCCCCTCGACAGCTCTCTGGGCACCCGGTCCCCCTCTGCCCCACCCCTC

946 GCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCTGCCTGGGCCCTGTTCAAGTGT

db
841 GCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCTGCCTGGGCGTGTTCAAGTGT

TCCACATAATACAGTATTCACACTCTTATCTTACAACTCCGCCACCGCCACTCT 1065
 TCCACATAATACAGTATTCACACTCTTATCTTACAACTCCGCCACCGCCACTCT 960
 TCTCAGTCTCCCAATCCCTGACCTTTGAGGCCCCCAGTGTATCTGACTCCCCC 1125
 TCTCAGTCTCCCAATCCCTGACCTTTGAGGCCCCCAGTGTATCTGACTCCCCC 1020
 TCCACAGCCCCCAGGCGATTTGTTCACTGTACTCTGTGGCAAGATGGGTCCAG 1185
 TCCACAGCCCCCAGGCGATTTGTTCACTGTACTCTGTGGCAAGATGGGTCCAG 1080
 ACCGACTTCAGGCACTAGAGGGGTGGACCTTGGGGGAGGAGCAAGAGACTG 1245
 ACCGACTTCAGGCACTAGAGGGGTGGACCTTGGGGGAGGAGCAAGAGACTG 1140
 TAGGCGAGAGTTCCTCAATGTAGGGGGGAGGAGCAAGAGTCTTCCCTTGA 1305
 TAGGCGAGAGTTCCTCAATGTAGGGGGGAGGAGCAAGAGTCTTCCCTTGA 1200
 TCCCTGTGATTTTAAACAGATATTTT 1341
 TCCCTGTGATTTTAAACAGATATTTT 1236

idard; DNA; 1030 BP.

(first entry)

NA.

is factor receptor; signal transducer molecule; TNF; APO4;
 abnormality; gestational abnormality; prostate cancer;
 APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 domain; immunogen; antibody preparation; breast carcinoma;
 man; ss.

Location/Qualifiers

1..627

/*tag= a

/product= "TNRL3"

98WO-US018393.

97US-00924634.

WASHINGTON.

191/17.

590.

is Factor family receptor polypeptides and ligands -
 agnosis and treatment of prostate cancer and developmental
 abnormalities.

Fig 13A; 156pp; English.

n describes isolated Tumor Necrosis Factor (TNF) family
 peptides: APO4, APO6, APO8 and APO9 or their active
 id isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or

their active fragments. APO4 is useful for diagnosing prostate c
 determining levels of APO4 in an individual. Prostate cancer can
 treated using APO4 selective binding agents linked to a therapeu
 moiety. APO4 polypeptides are also useful for identifying select
 binding agents, useful in diagnosis/treatment of disease by bind
 agents to the polypeptide/active fragment which is extracellular
 expressed on the cell surface. The binding is preferably perform
 vivo. APO4 polypeptides/ active fragments are also useful for sc
 for agonists and antagonists by binding and observing the change
 activity. Effective pharmacological agents useful in diagnosis o
 treatment of disease are also identified using APO4 polypeptides
 fragments and APO4 signal transducer molecules that specifically
 with a cytoplasmic domain of APO4 and detecting a change in leve
 activity. The method is performed in vivo or in vitro. APO polyp
 are all useful as immunogens for preparing antibodies. APO4 is a
 useful for diagnosis/treatment of developmental or gestational
 abnormalities. APO8 was transfected to human breast carcinoma ce
 MCF-7, and induced apoptosis

Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 U; 0 Other;

Query Match 60.7%; Score 833.4; DB 2; Length 1030;
 Best Local Similarity 99.9%; Pred. No. 1.1e-167;
 Matches 834; Conservative 0; Mismatches 1; Indels 0;

QY	229	GTCACTTTGGGAGCCGGGCATCGCTGTCGGCCAGGAGCCTGCCAGGAGGAGC
DB	1	GTCACTTTGGGAGCCGGGCATCGCTGTCGGCCAGGAGCCTGCCAGGAGGAGC
QY	289	GCAGAGGAGGACCCAGGACCGGTGCGAACTGAATCCCCAGACAGAAAGCCAGG
DB	61	GCAGAGGAGGACCCAGGACCGGTGCGAACTGAATCCCCAGACAGAAAGCCAGG
QY	349	GGCGCTTTCTGAAACCGACTAGTTCGGCTCGCAGAGTGACCTAAAGGCCGGA
DB	121	GGCGCTTTCTGAAACCGACTAGTTCGGCTCGCAGAGTGACCTAAAGGCCGGA
QY	409	CGGCTCTGAAGAGCGGATCGAGCCCATTTATGAAGTTTCATCCAGGACCTGGACAGG
DB	181	CGGCTCTGAAGAGCGGATCGAGCCCATTTATGAAGTTTCATCCAGGACCTGGACAGG
QY	469	GGCAGGCGAGGTGTGACGGGACAGTGTAGTGGCTGGGAGGAGCCAGAAATCAACAC
DB	241	GGCAGGCGAGGTGTGACGGGACAGTGTAGTGGCTGGGAGGAGCCAGAAATCAACAC
QY	529	AGCCCTCTGCGCTACAAACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCT
DB	301	AGCCCTCTGCGCTACTACCCGCGAGATCGGGAGTTTATAGTCACCCGGGCTGGGCT
QY	589	TACCTGTACTGTCAAGTGACCTTTTGTATGAGGGGAGGCTGTCTACCTGAAGCTGG
DB	361	TACCTGTACTGTCAAGTGACCTTTTGTATGAGGGGAGGCTGTCTACCTGAAGCTGG
QY	649	CTGGTGGATGTGTGCTGGCCCTGGGCTGCTGGAGGAAATTCAGCCACTGGCGC
DB	421	CTGGTGGATGTGTGCTGGCCCTGGGCTGCTGGAGGAAATTCAGCCACTGGCGC
QY	709	TCCCTCGGGGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTGGCCCTGGCGC
DB	481	TCCCTCGGGGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTGGCCCTGGCGC
QY	769	TCCCTCGGGGATCCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCCT
DB	541	TCCCTCGGGGATCCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCCT
QY	829	TACTTCGAGCTTTCAGGTTCACTGAGGGGCCCTGTGTCTCCCCACAGTCTCCCP
DB	601	TACTTCGAGCTTTCAGGTTCACTGAGGGGCCCTGTGTCTCCCCACAGTCTCCCP
QY	889	GCCGGCTCCCTCTGCACAGCTCTCTGGGCGACCCGGTCCCTCTGCCCCACCTCAGC
DB	661	GCCGGCTCCCTCTGCACAGCTCTCTGGGCGACCCGGTCCCTCTGCCCCACCTCAGC

us-09-245-198a-3.rng

Qy	106	ATGGCGCGCCGCTCGAGAGCCAGAGCGAGGGGCGCGGGGGGAGCCGGGGCACCG
Db	1	ATGGCGCGCCGCTCGAGAGCCAGAGCGAGGGGCGCGGGGGGAGCCGGGGCACCG
Qy	166	CTGGTCCCGCTCGCGCTGGGCGCTGGGCGCTGGGCGCTGGCGCTGGCGCTGGCT
Db	61	CTGGGCTCCCGCTCGAGAGCCAGAGCGAGGGGCGCGGGGGGAGCCGGGGCACCG

Db 121 GTGGTCAGCCTGGGAGCTGGGCAACGCTGTCTGCCAGGAGCCTTCTCAGGAGG

Db 181 ACAGCAGAGGACCGCCGGAGCCCCCTGAACTGAATCCCCAGACAGAGGAAAGCCF

QY
346 CCTGCGCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGAGTGCACCTAAAGGCCG

[illegible]

100

421 TCCAGCCCCTCTGGGCTACGACCGCCAGATTGGGAATTACAGTCATCAGGGCTGG

[illegible]

QY 386 TACACCTGTACTGT CAGGTGCACCTTTGATGAGGGGAGGCTGTCTACCTGAAGCT

Db 481 TACTACCTGTA CTGT CAGGTG CACTTTGATGAGGGAAGGCTGTCTACCTGAAGCT

Db
541 TTGCTGGTGAACGGTGTGCTGGCCCTGCGCTGCCCTGGAAGAAATTCTCAGCCACAGC

QY 706 AGTTCCCTCGGGCCCCAGCTCCGCCCTCTGCCAGGTGCTGGGCTGTTGGCCCTGCG

db
601 AGCTCTCCTGGGCCCCAGCTCCGTTTGTCGCCAGGTGCTGGGCTGTTGCCGCTGCG

QY 766 GGTCTCTCCCTGCGGATCCGCACCCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTT

Db 661 GGTCTTCCCTTCGGATCCGCACCCCTCCCTGGGCTCATCTTAGGCTGCCCCCTT

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

791

QY	941	CAGCGCTCTTTGCTCCAGACCTGCCCTCCCTCTAGAGGCTGCCCTGGGCGCTGTTC
Db	841	C-GCTGCTCTTGGTCCAGTCTGTCTCC--TCAAGGCGCGAGAGCTGTGTCTC
QY	1001	GTTTTCATCCCAATATAAGTATTTCCCACTCTTATCTTCAACTCTCCCCACCC
Db	898	GTTTCCATCC-----ACAGAGTATCTTGTCTTCTTAAACATCCCATCCAC
QY	1061	ACTCTCCACTCACTAGCTCCCAATCCGTAGCCCTTTGAGGCCCGCCAGTGATCTCC

Db 951 ACTATCCACCTCACTAGCTCCCAAGCCCTAC-----TATCCCTC

1121 CCCCCCTGGCCACAGACCCCCCAGGGCATTTGTGTTCACTGTA

996

DP 596 CCCCACCCACTCACCACCGACCGTGTATTGACTTTGTGTCAC-----

QY 1181 TCCAGAAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCAAGAGCCAA

-----CAGGCACTGAGATGGGCTGGACCTGGTGGCAGGAAGCCAGAGA 1082
TGGCCCTAGGCCAGGAGTCCCAATGTGAGGGCGAGA-AACAAGACAAGCTCCTC 1299
TGGGACTAGGCCAGGAGTCCCACTGTGAGGGGGAAGAGCTGGGGACAAGCTCCTC 1142
TGAGAAATCCCTGTGGGATTTTAAACAGAGATATTATTTTATTATTATTGTGACAAA 1359
GGA-----TCCCTGTGGATTTTGAAG--AGATACTATTATTATTATTATTATTGTGACAAA 1196
TTGA 1366
TTAA 1203

standard; DNA; 898 BP.

(first entry)

vector pDC409-LZ-TWEAK fusion protein-encoding DNA.

cellular domain; tumour necrosis factor; TNF; angiogenesis;
sclerolisation; diabetic retinopathy; neovascular glaucoma;
na; retinopathy of prematurity; retrolental fibroplasia;
itis; macular degeneration; arthritis; rheumatism; ds;
neovascularisation; psoriasis; metastatic condition;
tumor; sarcoma; carcinoma; benign tumor; haemophilic joint;
condition; myocardial angiogenesis; ischaemia; human;
vascular adhesion; telangiectasia; coronary atherosclerosis;
ic plaque neovascularisation; TWEAK receptor; TWEAKR;
therosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
n.

Location/Qualifiers

52..873

/tag= a

/product= "Fusion protein comprising a growth hormone
leader, a leucine zipper multimerisation domain, and
human TWEAK extracellular domain"

12.

2000WO-US034755.

99US-0172878P.

2000US-0203347P.

TEX CORP.

975/44.
-499.

giogenesis in a mammal for treating diseases mediated by
e.g. solid tumors and vascular deficiencies of cardiac or
ssue, by administering antagonist or agonist of TWEAK

ge 39-40; 46pp; English.

represents a DNA from the expression vector pDC409-LZ-TWEAK,
a fusion protein comprising a growth hormone leader, a

CC leucine zipper multimerisation domain, and the extracellular dom
CC human TWEAK. The fusion protein was used in the isolation of hum
CC receptor (TWEAKR)-expressing clones from a COS cell human CDNA l
CC The TWEAK protein is a member of the tumour necrosis factor (TNF
CC and induces angiogenesis. TWEAKR may therefore be used to screen
CC develop TWEAKR agonists and antagonists for the modulation of
CC angiogenesis, to be used in the treatment and diagnosis of human
CC The disorders mediated by angiogenesis include ocular disorders
CC characterised by ocular neovascularisation such as diabetic reti
CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity
CC retrolental fibroplasia, rubeosis, uveitis, macular degeneration
CC corneal graft neovascularisation, and inflammatory diseases such
CC arthritis, rheumatism and psoriasis. Other treatable diseases in
CC malignant and metastatic conditions such as sarcomas and carcino
CC benign tumours and preneoplastic conditions, myocardial angione
CC haemophilic joints, scleroderma, vascular adhesions, atheroscler
CC plaque neovascularisation, telangiectasia, wound granulation, co
CC atherosclerosis, peripheral atherosclerosis and ischaemia
XX
SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 U; 0 Other;

Query Match 45.8%; Score 629.2; DB 4; Length 898;
Best Local Similarity 99.5%; Pred. No. 3.1e-124;
Matches 631; Conservative 0; Mismatches 3; Indels 0;

QY 232 AGTTTGGGAGCGCGGATCGCTGTCGCGCCAGGAGCCCTGCCAGGAGAGCTGG
Db 250 AGTTTGGGAGCGCGGATCGCTGTCGCGCCAGGAGCCCTGCCAGGAGAGCTGG
QY 292 GAGGAGGACAGGACCCCTGCGAACTGAATCCCGAGAGAGAAAGCCAGGATC
Db 310 GAGGAGGACAGGACCCCTGCGAACTGAATCCCGAGAGAGAAAGCCAGGATC
QY 352 CTTTCTCCACCGGACTAGTTCGGCTCGCAGAGTGCACCTAAAGGCCGGA
Db 370 CTTTCTCCACCGGACTAGTTCGGCTCGCAGAGTGCACCTAAAGGCCGGA
QY 412 GCTCGAAGAGCGGATCGCAGCCCATTTATCAAGTTTCATCCAGACCTGGACAGGACG
Db 430 GCTCGAAGAGCGGATCGCAGCCCATTTATCAAGTTTCATCCAGACCTGGACAGGACG
QY 472 CAGGAGGTGTGGACGGACAGTGTGCTGGGAGGAGCCAGAAATCAACAGCT
Db 490 CAGGAGGTGTGGACGGACAGTGTGCTGGGAGGAGCCAGAAATCAACAGCT
QY 532 CCTCTGCGCTACACCGCCAGATCGGCGAGTTTATAGTCACCCGGCTGGGCTCT
Db 550 CCTCTGCGCTACACCGCCAGATCGGCGAGTTTATAGTCACCCGGCTGGGCTCT
QY 592 CTGTACTGTCAAGTGTCACTTTTGTATGAGGGGAGGCTGTCTACCTGAAGCTGGACT
Db 610 CTGTACTGTCAAGTGTCACTTTTGTATGAGGGGAGGCTGTCTACCTGAAGCTGGACT
QY 652 GTGATGTGTGTGGCCCTCGGCTGTGCGAGGTGTGGGCTGTGGCCCTGCGGCCAG
Db 670 GTGATGTGTGTGGCCCTCGGCTGTGCGAGGTGTGGGCTGTGGCCCTGCGGCCAG
QY 712 CTCGGGCCCCAGCTCCGCTCTGCGAGGTGTGGGCTGTGGCCCTGCGGCCAG
Db 730 CTCGGGCCCCAGCTCCGCTCTGCGAGGTGTGGGCTGTGGCCCTGCGGCCAG
QY 772 TCCCTGGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCTCA
Db 790 TCCCTGGGATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCTCA
QY 832 TTCGATCTTTCAGGTTACTGTAGGGGCCCTGG 865
Db 850 TTCGATCTTTCAGGTTACTGTAGGGGCCCTGG 883

RESULT 14

AAV18599

ID AAV18599 standard; cDNA; 1168 BP.

CTGGGCAACGGCTGTCGCCAGGAGCCTTCTCAGGAGGAGCTGACAGCGAGGAC 121

ndard; DNA; 701 BP.

(first entry)

DNA.

sis factor receptor; signal transducer molecule; TNF; APO4;
1 abnormality; gestational abnormality; prostate cancer;
APO4; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
domain; immunogen; antibody preparation; breast carcinoma;
ouse; ss.

Location/Qualifiers
1. 636
/*tag= a
/product= "TNRL3"

98WO-US018393.
97US-00924634.
WASHINGTON.

3191/17.
1591.

rosis Factor family receptor polypeptides and ligands -
agnosis and treatment of prostate cancer and developmental
al abnormalities.

Fig 13B; 156pp; English.

n describes isolated Tumor Necrosis Factor (TNF) family
peptides: APO4, APO6, APO8 and APO9 or their active
d isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
ragments. APO4 is useful for diagnosing prostate cancer by
evels of APO4 in an individual. Prostate cancer can also be
APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
s, useful in diagnosis/treatment of disease by binding of
polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
olypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the change in APO4
ective pharmacological agents useful in diagnosis or
disease are also identified using APO4 polypeptides/active
APO4 signal transducer molecules that specifically interact
asmic domain of APO4 and detecting a change in level of APO4
method is performed in vivo or in vitro. APO polypeptides
d as immunogens for preparing antibodies. APO4 is also
agnosis/treatment of developmental or gestational
APO8 was transfected to human breast carcinoma cell line
duced apoptosis

BP; 139 A; 210 C; 203 G; 149 T; 0 U; 0 Other;
37.8%; Score 519.2; DB 2; Length 701;
arity 87.3%; Pred. No. 7.3e-101;
onservative 0; Mismatches 83; Indels 0; Gaps 0;

220 CTGGCCGTGGTCTAGTTTGGGAGCCGGGAGTCGCTGTCGCCGCCAGGAGCGTGGCC
1 CTGGTGTGGTCTAGCTTGGGAGCTGGGCAACGCTGTCTGCCAGGAGCGCTTCTC
280 GAGCTGGTGGCAGAGGAGGACAGGACCCGTCGGAACCTGAATCCCGCAGACAGAG
61 GAGCTGACAGCAGAGGAGCCCGGGAGCCCGCTGAATGAATCCCGCAGACAGAG
340 CAGGATCCCTGGCCCTTCTGTAACCGACTAGTTCGGCCCTCGCAGAGAGTGCACCTA
121 CAGGATGGTACCTTCTTGGAACTAGTTCGGCCCTCGAAGAGAGTGTCTCTA
400 CGGAAACACGGGCTCGAAGAGCGATCGAGCCCATTTATGAGTTTCATCCAGAC
181 CGGAGGCGGGCCCTCGCGAGCTATTGACGCCCATTTATGAGTTTCATCCTCGGC
460 CAGGACGGAGCGCAGCGAGTGTGGAACGGGACAGTGAAGTGGGAGGAGGACCA
241 CAGGATGGAGCACAGCAGGTGTGGATGGACAGTGAAGTGGGAGGAGGACCA
520 AACAGCTCCAGCCCTCTGGCTACACCGCCAGATCGGGAGTGTATAGTCAACC
301 AACAGCTCCAGCCCTCTGGCTACACCGCCAGATCGGGAGTGTATAGTCAACC
580 GGGCTCTACTACTGTACTGTCTGAGTGTGATGAGGGGAGGAGTGTCTTACC
361 GGGCTCTACTACTGTACTGTCTGAGTGTGATGAGGGGAGGAGTGTCTTACC
640 CTGGAATCTGCTGGTGGATGGTGTCTGGCCCTCGCTGCTGGAGGAAATTTCTCAG
421 CTGGAATCTGCTGGTGGATGGTGTCTGGCCCTCGCTGCTGGAGGAAATTTCTCAG
700 GGGCCAGTTCCTCGGGCCCCAGCTCGCCCTCTGCGAGGTGTCTGGGCTGTGG
481 GAGCAAGCTCTCTCGGGCCCCAGCTCGCTTGTGTCAGGTGTCTGGGCTGTGG
760 CGGCCAGGCTCTCTCGGATCCGACCCCTCGGCCCATCTCAAGGCTG
541 CGGCCAGGCTCTCTCGGATCCGACCCCTCGGCCCATCTCAAGGCTG
820 TTCCTCAGCTTCTCGGATCTTCCAGGTTCAGTGGGGGCGCTGTCTCC 87:
601 TTCCTCAGCTTCTCGGATCTTTCAGGTTCAGTGGGGGCGCTGTCTCTCC 65:

Search completed: April 7, 2004, 21:32:21
Job time : 528.669 secs

06:25:17 2004

us-09-245-198a-3.rnpb

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

ic search, using sw model

ril 7, 2004, 23:15:35 ; Search time 523.048 Seconds
(without alignments)
9837.847 Million cell updates/sec

-09-245-198A-3

73

atgcattgttagacttga.....gacaaaattgtgataaatgg 1373

ENTITY_NUC

pop 10.0 , Gapext 1.0

70632 seqs, 1873875610 residues

ts satisfying chosen parameters: 4941264

gth: 0

gth: 2000000000

inimum Match 0%

aximum Match 100%

isting first 45 summaries

ublished Applications NA:*

/cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:
/cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:
/cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:
/cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:
/cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
/cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:
/cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:
/cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:
/cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
ad by analysis of the total score distribution.

SUMMARIES

ch	Length	DB	ID	Description
1.0	1364	9	US-09-822-849A-19	Sequence 19, Appl
1.2	1353	14	US-10-210-951-3	Sequence 3, Appl
1.2	1353	14	US-10-211-884-3	Sequence 3, Appl
1.6	1306	12	US-10-202-062-23	Sequence 23, Appl
1.6	1306	14	US-10-272-411-16	Sequence 16, Appl
1.6	1306	14	US-10-218-547-23	Sequence 23, Appl
1.6	1306	14	US-10-272-328A-16	Sequence 16, Appl
1.6	1306	14	US-10-310-793-29	Sequence 29, Appl
1.8	898	9	US-09-742-454A-1	Sequence 1, Appl
1.8	898	9	US-09-883-777-1	Sequence 1, Appl
1.4	493	10	US-09-918-995-21225	Sequence 21225, A
1.5	408	9	US-09-960-352-2197	Sequence 2197, Ap
1.7	213	12	US-10-085-783A-55176	Sequence 55176, A
1.7	213	15	US-10-242-535A-55176	Sequence 55176, A
1.7	264	9	US-09-983-965-2183	Sequence 2183, Ap

C 16	67.2	4.9	3163	14	US-10-017-161-1857	Sequence
C 17	67.2	4.9	3163	15	US-10-292-798-1513	Sequence
C 18	65.4	4.8	1064	9	US-09-804-682-29	Sequence
C 19	63.8	4.6	5452	14	US-10-017-161-1481	Sequence
C 20	63.8	4.6	5452	15	US-10-292-798-1189	Sequence
C 21	63.4	4.6	778	12	US-10-424-599-54839	Sequence
C 22	62.2	4.5	12733	14	US-10-032-393-47	Sequence
C 23	62.2	4.5	12733	14	US-10-032-393-8	Sequence
C 24	62	4.5	815	12	US-10-424-599-20495	Sequence
C 25	61.4	4.5	1117	14	US-10-017-161-1403	Sequence
C 26	61.4	4.5	1117	15	US-10-292-798-1141	Sequence
C 27	60	4.4	60	10	US-09-908-975-13797	Sequence
C 28	60	4.4	60	10	US-09-908-975-13139	Sequence
C 29	60	4.4	60	10	US-09-908-975-31341	Sequence
C 30	60	4.4	60	10	US-09-908-975-31596	Sequence
C 31	60	4.4	60	10	US-09-908-975-31597	Sequence
C 32	59.6	4.3	3133	14	US-10-017-161-1483	Sequence
C 33	59.6	4.3	3133	15	US-10-292-798-1191	Sequence
C 34	56	4.1	1065	9	US-09-804-682-33	Sequence
C 35	55.6	4.0	771	12	US-10-424-599-107958	Sequence
C 36	55.2	4.0	580	12	US-10-424-599-2320	Sequence
C 37	55	4.0	791	12	US-10-424-599-64474	Sequence
C 38	54.4	4.0	1327	15	US-10-398-221-1265	Sequence
C 39	54	3.9	931	14	US-10-198-846-8585	Sequence
C 40	53.2	3.9	53522	10	US-09-904-968A-1	Sequence
C 41	52.6	3.8	636	12	US-10-424-599-72906	Sequence
C 42	52.2	3.8	594	12	US-10-142-426-10	Sequence
C 43	52.2	3.8	594	14	US-10-123-155-10	Sequence
C 44	52.2	3.8	594	14	US-10-146-731-10	Sequence
C 45	52.2	3.8	594	14	US-10-140-472-10	Sequence

ALIGNMENTS

RESULT 1

US-09-822-849A-19
; Sequence 19, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEIN
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822, 849A
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 1364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-19

Query Match 98.0%; Score 1345.8; DB 9; Length 1364;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1347; Conservative 0; Mismatches 2; Indels 0; G;
QY 25 TCCTCCCGCCGCTCCCTCCCTCCCGATCCCTCCGCTCCCGGATCGGGGCGGCT
Db 5 TCCTCCCGCCGCTCCCTCCCTCCCGATCCCTCCGCTCCCGGATCGGGGCGGCT
QY 85 CAGGCACAGCCCCCGCCCGCCCGATGCGCGCTCCGAGCCAGAGGCGGAGGCGGCGC

[illegible]

US-10-21

Query Match 96.2%; Score 1320.2; DB 14; Length 1353;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1322; Conservative 0; Mismatches 3; Indels 0; C

49 CGATCCCTCGGGTCCCGGATCGGGGGGGCGGTGAGCGAGGCACAGCCCCCGGCCCG
1 CGATCCCTCGGGTCCCGGATCGGGGGGGCGGTGAGCGAGGCACAGCCCCCGGCCCG
109 GCAGCCCGTCCGAGCCAGAGGCGAGGGGGGGCCCGGGGGAGCGCGGGCAGCGCC

Db	1141	ACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGACGGAAGCCAAAGAGACCT
Qy	1249	CTAGGCCAGGAGTTTCCCAAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCTCTTC
Db	1201	CTAGGCCAGGAGTTTCCCAAAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCTTC
Qy	1309	TTCCCTGTGGATTTTTAAACACAGATATTATTTTTATTATTGTGACAAAATGT
Db	1261	TTCCCTGTGGATTTTTAAACACAGATATTATTTTTATTATTGTGACAAAATGT
Qy	1369	AATGG 1373
Db	1321	AATGG 1325

US-10-211-884-3

```

Query Match      96.2%; Score 1320.2; DB 14; Length 1953;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 132; Conservative 3; Indels 0; Gaps
0;

49  CGATCCTCGGTCCTCCGGATGGGGGGGCGGTGAGGACGACAGCCCCCGCCCC
1  CGATCCTCGGTCCTCCGGATGGGGGGGCGGTGAGGACGACAGCCCCCGCCCC

```


3CCGTCGAGCCAGAGCGGCGCGCGCGGCGAGCCGCGCCCTGCTG 168
3CCGTCGAGCCAGAGCGGCGCGCGGCGAGCCGCGCCCTGCTG 120
CGCTCGCGCTGGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 228
CGCTCGCGCTGGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 180
AGTTGGGAGCGCGGCGATCGCTGCTCGCGCCAGGAGCTGCCAGGAGGAGCTGGTG 288
AGTTGGGAGCGCGGCGATCGCTGCTCGCGCCAGGAGCTGCCAGGAGGAGCTGGTG 240
3AGGAGCAGAGACCGCGTCGAACTGAATGCCCGCAGACAGAAAGCCAGGATCCCT 348
3AGGAGCAGAGACCGCGTCGAACTGAATGCCCGCAGACAGAAAGCCAGGATCCCT 300
CCTTTCTGAACCGACTAGTTCCGCTCGCAGAGTGCACCTAAAGCCCGGAAACA 408
CCTTTCTGAACCGACTAGTTCCGCTCGCAGAGTGCACCTAAAGCCCGGAAACA 360
3CTCGAAGCGGATCGCAGCCCATTAATGAAGTTTCATCCAGACTGGACAGGACGGA 468
3CTCGAAGCGGATCGCAGCCCATTAATGAAGTTTCATCCAGACTGGACAGGACGGA 420
3AGGAGGTTGACCGGACAGTGAAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGTCC 528
3AGGAGGTTGACCGGACAGTGAAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGTCC 480
3CTCTGGCTACAAACCGCCAGATCGGGAGTTTATAGTACACCGGGCTGGGCTCTAC 588
3CTCTGGCTACAAACCGCCAGATCGGGAGTTTATAGTACACCGGGCTGGGCTCTAC 540
3TGTACTGTGAGTGACATTTGATGAGGGGAGGCTGTCTACCTGAAGTGGAGCTTG 648
3TGTACTGTGAGTGACATTTGATGAGGGGAGGCTGTCTACCTGAAGTGGAGCTTG 600
3TGGATGGTGTGCTGGCCCTGCGCTGGAGGAAATCTCAGGCACTGGGCGAGT 708
3TGGATGGTGTGCTGGCCCTGCGCTGGAGGAAATCTCAGGCACTGGGCGAGT 660
3TGGGCCCCCAGCTCTCGCAGTGTCTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 768
3TGGGCCCCCAGCTCTCGCAGTGTCTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 720
3CCTCGGATCCGACCCCTCCCTCGGCGCCATCTCAAGGCTGCCCCCTTCTCTACC 828
3CCTCGGATCCGACCCCTCCCTCGGCGCCATCTCAAGGCTGCCCCCTTCTCTACC 780
3TGGAGTCTTCCAGGTTCACTGAGGGGCGCTGGTCTGCCCAAGTCTGCCAGGCT 888
3TGGAGTCTTCCAGGTTCACTGAGGGGCGCTGGTCTGCCCAAGTCTGCCAGGCT 840
3GCTCCCTCGACAGCTCTCTGGGCGCCCGCTCCCTCTGCGCCACCCCTCAGCGCT 948
3GCTCCCTCGACAGCTCTCTGGGCGCCCGCTCCCTCTGCGCCACCCCTCAGCGCT 900
3GCTCGACAGCTCCCTCTAGAGGCTGCTGGGCTGTTCAGGTTTTCCTCA 960
3GCTCGACAGCTCCCTCTAGAGGCTGCTGGGCTGTTCAGGTTTTCCTCA 1008
3ACATAAATACAGTATCCCACTCTTATCTTACACTCCCGCCACCGCCCACTCTCA 1068
3ACATAAATACAGTATCCCACTCTTATCTTACACTCCCGCCACCGCCCACTCTCA 1020
3ACTAGTCCCCCAATCCCTGACCTTTGAGGCGCCCGCAGTGATCTGACTCCCCCTG 1128
3ACTAGTCCCCCAATCCCTGACCTTTGAGGCGCCCGCAGTGATCTGACTCCCCCTG 1080
3ACAGCCCCCAGGCGATTTGTTTCACTGTACTCTGTGGCAAGGATGGGTCCAGAAG 1188
3ACAGCCCCCAGGCGATTTGTTTCACTGTACTCTGTGGCAAGGATGGGTCCAGAAG 1140
3CACTTCAGGCACTAAGAGGGGCTGGACCTGGGCGGAGGAGCCAAAGAGACTGGGC 1248

1141 ACCCCACTTCAGCACTAAGAGGGGCTGCACCTGGCGGCGAGGAGCCAAAGAGAC
1249 CTAGGCGCAGAGTTCCCAAAATGTAGGGGCGGAGAAACAAGCAAGCTCTCTCCCTT
1201 CTAGGCGCAGAGTTCCCAAAATGTAGGGGCGGAGAAACAAGCAAGCTCTCTCCCTT
1309 TTCCCTGTGATTTTAAACACAGATATTTTATTATTATTATTTGACAAATGT
1261 TTCCCTGTGATTTTAAACACAGATATTTTATTATTATTATTATTGACAAATGT
1369 AATGG 1373
1321 AATGG 1325

RESULT 4

US-10-202-062-23
; Sequence 23, Application US/10202062
; Publication No. US20040038349A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.,
; TITLE OF INVENTION: Heteromultimeric TNF Ligand Family members
; FILE REFERENCE: PF559
; CURRENT APPLICATION NUMBER: US/10/202,062
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,838
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1306
; TYPE: DNA
; ORGANISM: human
US-10-202-062-23

Query Match 93.6%; Score 1285; DB 12; Length 1306;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1285; Conservative 0; Mismatches 0; Indels 0;

QY 89 CACAGCCCCCGCCCGCCATGGCCCGCTCGAGGCCAGAGGCGGAGGGGCGCGG
DB 1 CACAGCCCCCGCCCGCCATGGCCCGCTCGAGGCCAGAGGCGGAGGGGCGCGG
QY 149 AGCCGGGACACCGCTGCTGGTCCGCTCGGCTGGGCTGGGCTGGGCTGGG
DB 61 AGCCGGGACACCGCTGCTGGTCCGCTCGGCTGGGCTGGGCTGGGCTGGG
QY 209 TCGGCTCTCTGCTGGCCGCTGCTAGTTTGGGAGCGGCGCATCGTGTCCGCCA
DB 121 TCGGCTCTCTGCTGGCCGCTGCTAGTTTGGGAGCGGCGCATCGTGTCCGCCA
QY 269 CTGCCCCAGAGGAGCTGGTGGCAGAGGAGGACCCAGGACCCGTCGGAATCC
DB 181 CTGCCCCAGAGGAGCTGGTGGCAGAGGAGGACCCAGGACCCGTCGGAATCC
QY 329 CAGAGAAAGCCAGGATCCTGCGCTTTCTGAAACCGACTAGTTCGGCTCGCAG
DB 241 CAGAGAAAGCCAGGATCCTGCGCTTTCTGAAACCGACTAGTTCGGCTCGCAG
QY 389 CACCTAAAGGCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTTAAGT
DB 301 CACCTAAAGGCGGAAAAACACGGGCTCGAAGAGCGATCGCAGCCCATTTAAGT
QY 449 CAGGACCTGGACAGGACCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGT
DB 361 CAGGACCTGGACAGGACCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGT
QY 509 AAGCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATCGGGAGT
DB 421 AAGCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATCGGGAGT
QY 569 TCACCCGGGCTGGGCTCTACTACCTGTACTGTAGGTGCACTTTGTAGGGGAA

.CCGGCTGGGCTCTACTACCTGTAAGTGTGATGAGGGAAGCTG 540
 CCTGAAGCTGACTTTGCTGGTGGATGGTGTGTGGCCCTGCCTGCTGGAGGAAT 688
 CCTGAAGCTGGAATCTGCTGGTGGATGGTGTGTGGCCCTGGCTGGAGGAAT 600
 AGCCACTCGCGCCAGTTCCTCTGGCGCCCGCAGCTCCGCTCTGCCAGGTGTGGGC 748
 AGCCACTCGCGCCAGTTCCTCTGGCGCCCGCAGCTCCGCTCTGCCAGGTGTGGGC 660
 GGCCCTGGCGCCAGGGTCTCTCCCTGGGATCCGACCCCTCCCTGGCGCCATCTCA 808
 GGCCCTGGCGCCAGGGTCTCTCCCTGGGATCCGACCCCTCCCTGGCGCCATCTCA 720
 TGCCCTCTCTCACTACTTCCGACTCTCTCCAGGTTCACTAGGGGCGCTGTCT 868
 TGCCCTCTCTCACTACTTCCGACTCTCTCCAGGTTCACTAGGGGCGCTGTCT 780
 ACAGTGTCCAGGTCGGGTCCTCCCTCGACAGTCTCTGGGCACTCCGCTCCCT 928
 ACAGTGTCCAGGTCGGGTCCTCCCTCGACAGTCTCTGGGCACTCCGCTCCCT 840
 CCACCCCTCAGCGCTCTTGTCTCAGACCTGCCCTCCCTCTAGAGGTCGCTGG 988
 CCACCCCTCAGCGCTCTTGTCTCAGACCTGCCCTCCCTCTAGAGGTCGCTGG 900
 GTTCACTGTGTTTCCATCCCACTAATAATACAGTATCCCACTCTTATCTTCAACT 1048
 GTTCACTGTGTTTCCATCCCACTAATAATACAGTATCCCACTCTTATCTTCAACT 960
 CACGCGCACTCTCCACTACTAGCTCCCCAACTCCCTGACCCCTTGAGGCCCCCA 1108
 CACGCGCACTCTCCACTACTAGCTCCCCAACTCCCTGACCCCTTGAGGCCCCCA 1020
 TCTCGACTCCGCTCGGCCACAGACCCCGAGGCAATGTGTTCACTGTACTCTGTG 1168
 TCTCGACTCCGCTCGGCCACAGACCCCGAGGCAATGTGTTCACTGTACTCTGTG 1080
 AGATGGTCCAGAAAGCCCCACTTCAGGCACTAAGAGGGCTGCACTGGCGGCA 1228
 AGATGGTCCAGAAAGCCCCACTTCAGGCACTAAGAGGGCTGCACTGGCGGCA 1140
 CCCAAAGAGACTGGGCTTAGCGCAGGAGTTCCTCCAACTGTAGGGCGGAGAACAG 1288
 CCCAAAGAGACTGGGCTTAGCGCAGGAGTTCCTCCAACTGTAGGGCGGAGAACAG 1200
 CTCCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATATTTTATTTATT 1348
 CTCCTCCCTTGAGAAATTCCTGTGGATTTTAAACAGATATATTTTATTTATT 1260
 TGAATAAATGTTGATAAATGG 1373
 TGAATAAATGTTGATAAATGG 1285

```

Location US/10272411
JS20030100068A1
[ON:
as Jewish Hospital
Jonathan
), F. Patrick
elbaum, Steven
ON: RANKI MIMICS AND USES THEREOF
60019620-0202
ION NUMBER: US/10/272,411
DATE: 2002-10-15
ON NUMBER: 60/329,393
E: 2001-10-15
) NOS: 52
In version 3.1

```

;	SEQ ID NO 16
;	LENGTH: 1306
;	TYPE: DNA
;	ORGANISM: Homo sapiens
;	PUBLICATION INFORMATION:
;	DATABASE ACCESSION NUMBER: NCBI/AF030099.1
;	DATABASE ENTRY DATE: 1997-12-20
;	RELEVANT RESIDUES: (1)..(1306)
;	PUBLICATION INFORMATION:
;	DATABASE ACCESSION NUMBER: NCBI/NM_003809.2
;	DATABASE ENTRY DATE: 2002-10-07
;	RELEVANT RESIDUES: (1)..(1306)
;	PUBLICATION INFORMATION:
;	DATABASE ACCESSION NUMBER: NCBI/AF055872.1
;	DATABASE ENTRY DATE: 1998-05-04
;	RELEVANT RESIDUES: (1)..(1306)
;	PUBLICATION INFORMATION:
;	DATABASE ACCESSION NUMBER: NCBI/BC019047.1
;	DATABASE ENTRY DATE: 2001-12-11
;	RELEVANT RESIDUES: (1)..(1306)
;	PUBLICATION INFORMATION:
;	DATABASE ACCESSION NUMBER: NCBI/AF030100.1
;	DATABASE ENTRY DATE: 1997-12-20
;	RELEVANT RESIDUES: (1)..(1306)
US-	10-272-411-16
Query Match	93.6%; Score 1285; DB 14; Length 1306;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1285; Conservative	0; Mismatches 0; Indels 0; G
Qy	89 CACAGCCCCCGCCCCCATGGCCCGCTGCGAGCCAGAGCGCAGGGGGCGCCGC
Db	1 CACAGCCCCCGCCCCCATGGCCCGCTGCGAGCCAGAGCGCAGGGGGCGCCGC
Qy	149 AGCCGGGCACGCCCTCTGCTGGTTCGGTTCGGCTGGGCCTGGGCCTGGCGCTGGCC
Db	61 AGCCGGGCACGCCCTCTGCTGGTTCGGTTCGGTTCGGGCCTGGGCCTGGCGCTGGCC
Qy	209 TCGGCCTCTCTGCTGGCGGTGGTCAGTTTGCGGAGCCGGGCATTCGCTGTCGCCCAG
Db	121 TCGGCCTCTCTGCTGGCGGTGGTCAGTTTGCGGAGCCGGGCATTCGCTGTCGCCCAG
Qy	269 CTGCCACGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCCGTCGGAACTGAATCCC
Db	181 CTGCCACGAGGAGCTGGTGGCAGAGGAGGACCAGGACCCCGTCGGAACTGAATCCC
Qy	329 CAGAGAAGCCAGGATCTCGCGCTTCTTCTGAAACGACTAGTTTCGGCTCGCGACA
Db	241 CAGAAGAAGCCAGGATCTCGCGCTTCTTCTGAAACCGACTAGTTTCGGCTCGCGACA
Qy	389 CACCTAAGGCCCGGMAAACACGGGCTCGAAGAGCGATCGCAGGCCCATATATGAAGTT
Db	301 CACCTAAGGCCCGGMAAACACGGGCTCGAAGAGCGATCGCAGGCCCATATATGAAGTT
Qy	449 CAGCACTGGACAGGACGGAGCGCAGGCTGTGGACGGGACAGTAGTGAGTGGCTGG
Db	361 CACGACCTGGACAGGACGGAGCGCAGGACAGTGTGGACGGGACAGTAGTGAGTGG
Qy	509 AAGCCAGAACTAACAGCTTCAGCCCTCTCGCGCTTAACAACCGCCAGATCGGGAGATT
Db	421 AAGCCAGAAATCAACAGCTTCAGCCCTCTCGCGCTTAACAACCGCCAGATCGGGAGATT
Qy	569 TCACCCGGGCTGGGCTCTACTACTGTACTCTCAGGTGCACTTTGATGAGGGGAAG
Db	481 TCACCCGGGCTGGGCTCTACTACTGTACTCTCAGGTGCACTTTGATGAGGGGAAG
Qy	629 TCTACTGAAAGCTGGACTTGTCTGGTGGATGCTGTGCTGGGCCCTCGCGCTGCCTGGAG
Db	541 TCTACTGAAAGCTGGACTTGTCTGGTGGATGCTGTGCTGGGCCCTCGCGCTGCCTGGAG
Qy	689 TCTCAGCCAATCGCGCCAGTTCCCTCTCGGGCCCCAGCTCCGCTCTGCGAGGTGTCT
b	601 TCTCAGCCAATCGCGCCAGTTCCCTCTCGGGCCCCAGCTCCGCTCTGCGAGGTGTCT

de

Identity 100.0%; Pred. NO. 0;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CGGCAACCGCCCTGCTGGTCCCGCTCGCGCTGGGCCCTGGCGCTGGCCCTGCC 208

3CCTCCTGCTGGCCGTGGTCAGTTTGGGGAGCCGGGCATCGCTGTCGCCCCAGGAGC 268

CCAGGAGGAGCTGGTGCAGAGGAGGACCCAGACCGTCCGGAACCTGAATCCCCAGA 328

AGAAAGCCAGGATCCTGTGGCCTTTCTTGAAACCGACTAGTTCGGCCTCGCAGAAAGTG 388

TTAAAGGCCGGAAAAACACGGGGCTCGAAGAGCGGATCGCAGCCCCATTATGAAGTTCATC 448

1ACTTGGACAGGACGGAGCGCAGGCAGGTGTGGACGGACAGTGAGTGGGAGG 508

5' CAGAAATCAACAGCTCCAGCCCTCTGGCTACAACCGCCAGATCGGGGAGTTTATAG 568

'CCGGGCTGGGCTCTACTACCTGTACTGTCAGGTGCACCTTTGATGAGGGGAAGGCTG 628

CCTGAAGCTGGACTTGCTGGTGGATGCTGTCTGGCCCTGCGCTGCCTGGAGGAAT 688

RESULT 9

; Sequence 1, Application US/09742454A

; GENERAL INFORMATION:

FILE OF INVENTION: TWEAK Receptor
FILE REFERENCE: 2069-B

; CURRENT FILING DATE: 2000-12-19

; PRIOR FILING DATE: 1999-12-20

PRIOR FILING DATE: 2000-05-10

```

; SOFTWARE: PATECIII VER. 2.0
: SEQ ID NO 1

```

TYPE: DNA

; FEATURE:

; LOCATION: (52) .. (873)

06:25:17 2004

us-09-245-198a-3.rnpb

ION: Description of Artificial Sequence: human TWEAK
ION: fusion protein construct

45.8%; Score 629.2; DB 9; Length 898;
arity 99.5%; Pred. No. 1.1e-162;
onservative 0; Mismatches 3; Indels 0;
TGGGAGCGGGCATCGCTGTCGCGCCAGGAGCTGCCAGGAGAGCTGGTGGCA 291
TGGGAGCGGGCATCGCTGTCGCGCCAGGAGCTGCCAGGAGAGCTGGTGGCA 309
AGGACCGAGCCCTCGGAACTGATCCCGACAGAGAAAGCCAGGATCCTCGG 351
AGGACCGAGCCCTCGGAACTGATCCCGACAGAGAAAGCCAGGATCCTCGG 369
TCCTGAACCCGACTAGTTCCGCTCGCAGAGTGCACCTAAAGCGCGGAAACACGG 411
TCCTGAACCCGACTAGTTCCGCTCGCAGAGTGCACCTAAAGCGCGGAAACACGG 429
GAAGAGGATCGGAGCCCATTTATGAAGTTTCATCCAGCACTGGACAGGACGGAGCG 471
GAAGAGGATCGGAGCCCATTTATGAAGTTTCATCCAGCACTGGACAGGACGGAGCG 489
CAGGTGCGGAGGACAGTGAAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC 531
CAGGTGCGGAGGACAGTGAAGTGGCTGGGAGGAGCCAGAAATCAACAGCTCCAGC 549
TGGCTACAAACCGCAGATCGGGAGTTTATAGTCACCCGCGCTGGCTCTACTAC 591
TGGCTACAAACCGCAGATCGGGAGTTTATAGTCACCCGCGCTGGCTCTACTAC 609
ACTGTCAAGTGCATTTGATGAGGGGAGGCTGTCTACCTGAAGCTGGACTTGGCTG 651
ACTGTCAAGTGCATTTGATGAGGGGAGGCTGTCTACCTGAAGCTGGACTTGGCTG 669
ATGCTGCTGCGCCCTCGCTGCTGAGGAAATCTCAGGCACTGGCGCCAGTTCC 711
ATGCTGCTGCGCCCTCGCTGCTGAGGAAATCTCAGGCACTGGCGCCAGTTCC 729
TGGCCAGCTCCGCTCTGCGAGGTGCTGGGCTGTTGGCCCTGGCGCCAGGTTCC 771
TGGCCAGCTCCGCTCTGCGAGGTGCTGGGCTGTTGGCCCTGGCGCCAGGTTCC 789
TGGGATCCGCACTCCCTCGGCGCCATCTCAAGGCTGCCCCCTTCTCCTCCTAC 831
TGGGATCCGCACTCCCTCGGCGCCATCTCAAGGCTGCCCCCTTCTCCTCCTAC 849
TACTCTCCAGGTTCACTGAGGGGCGCTGG 865
TACTCTCCAGGTTCACTGAGGGGCGCG 883

cation US/09883777
10110853A1

ON: Steven R.
ON: TWEAK RECEPTOR

2968-C
ION NUMBER: US/09/883,777
DATE: 2001-06-18
N NUMBER: US 60/172,878
E: 1999-12-20
N NUMBER: US 60/203,347
E: 2000-05-10
N NUMBER: PCT/US00/34755
E: 2000-12-19
N NUMBER: US 09/742,454
E: 2000-12-19
NOS: 16
In version 3.1

; SEQ ID NO 1
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: TWEAK fusion protein construct
; NAME/KEY: CDS
; LOCATION: (52)..(873)
; OTHER INFORMATION:
US-09-883-777-1

Query Match 45.8%; Score 629.2; DB 9; Length 898;
Best Local Similarity 99.5%; Pred. No. 1.1e-162;
Matches 631; Conservative 0; Mismatches 3; Indels 0; C

QY 232 AGTTTGGGGAGCCGGGATCGCTGTCGCGCCAGGAGCTGCCAGGAGAGCTGGT
Db 250 AGTTTGGGGAGCCGGGATCGCTGTCGCGCCAGGAGCTGCCAGGAGAGCTGGT
QY 292 GAGGAGGACCAAGGACCCGCTCGGAACTGAATCCCGACAGAGAAAGCCAGGATCC
Db 310 GAGGAGGACCAAGGACCCGCTCGGAACTGAATCCCGACAGAGAAAGCCAGGATCC
QY 352 CTTTCTCTGAACCGACTAGTTTGGCTCGCAGAGTGCACCTAAAGCGCGGAAAC
Db 370 CTTTCTCTGAACCGACTAGTTTGGCTCGCAGAGTGCACCTAAAGCGCGGAAAC
QY 412 GCTGGAAGAGGATCGCAGGCCATTTATGAAGTTTCATCCAGCACTGGACAGGACGG
Db 430 GCTGGAAGAGGATCGCAGGCCATTTATGAAGTTTCATCCAGCACTGGACAGGACGG
QY 472 CAGCAGGCTGTGACGCGGACAGTGAAGTGGCTGGGAGGAGCCAGAAATCAACAGCTC
Db 490 CAGCAGGCTGTGACGCGGACAGTGAAGTGGCTGGGAGGAGCCAGAAATCAACAGCTC
QY 532 CTTCTGCGCTTACAAACCGCAGATCGGGAGTTTATAGTCACCCGCGCTGGCTCTTA
Db 550 CTTCTGCGCTTACAAACCGCAGATCGGGAGTTTATAGTCACCCGCGCTGGCTCTTA
QY 592 CTGTACTGTCAAGTGCATTTGATGAGGGGAGGCTGTCTACCTGAAGCTGGACTT
Db 610 CTGTACTGTCAAGTGCATTTGATGAGGGGAGGCTGTCTACCTGAAGCTGGACTT
QY 652 GTGGATGTTGCTGTCGCTCGCTGCTGAGGAGGAAATTTCTCAGCCACTGGCGCCAG
Db 670 GTGGATGTTGCTGTCGCTCGCTGCTGAGGAGGAAATTTCTCAGCCACTGGCGCCAG
QY 712 CTCGGGCCCCAGCTCCGCTCTGCGAGGTGCTGCGGCTGTTGGCCCTGGCGCCAGG
Db 730 CTCGGGCCCCAGCTCCGCTCTGCGAGGTGCTGCGGCTGTTGGCCCTGGCGCCAGG
QY 772 TCCTTCGGATCGGACCCCTCCCTCGGCGCCATCTCAAGGCTGCCCTTCTCCTCAC
Db 790 TCCTTCGGATCGGACCCCTCCCTCGGCGCCATCTCAAGGCTGCCCTTCTCCTCAC
QY 832 TTCGAGTCTTCCAGGTTCACTGAGGGGCGCTGG 865
Db 850 TTCGAGTCTTCCAGGTTCACTGAGGGGCGCG 883

RESULT 11

US-09-918-995-21225
; Sequence 21225, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20

ID NOS: 38054
SEQ for Windows Version 3.0

io sapiens

c_feature

... (493)
TION: n = A, T, C or G
25

32.4%; Score 445.4; DB 10; Length 493;

larity 99.8%; Pred. No. 3.1e-112;
Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ACTCTTCCAGTTCACTGAGGGGCTTGTCTCCCCACAGTCGTCCTGAGGCTGCGG 893

ACTCTTCCAGTTCACTGAGGGGCTTGTCTCCCCACAGTCGTCCTGAGGCTGCGG 106

CCCTCGACAGCTCTCTGGGACACCGGTCCTCTGCCCCCAGCTGCTGAGGCTGCGG 953

CCCTCGACAGCTCTCTGGGACACCGGTCCTCTGCCCCCAGCTGCTGAGGCTGCGG 166

CAGACCTGCCCCCTCTCTAGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCGG 1013

CAGACCTGCCCCCTCTCTAGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCGG 226

AAATACAGTATCCCACTCTTATCTTAACTTCCCACTGCTGAGGCTGCTGAGGCTGCGG 1073

AAATACAGTATCCCACTCTTATCTTAACTTCCCACTGCTGAGGCTGCTGAGGCTGCGG 286

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 1133

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 346

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 1193

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 406

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 1253

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 466

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 1280

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG 493

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

CTCTCCCAATCCCTGACCTTTGAGGCCCCAGTGTCTGCTGAGGCTGCTGAGGCTGCGG

|||||
ATTCCTGTGGATTTTAAACAGATATTATTTATTATTATTGTGACAAATGT 180
|||||
AAATGG 1373
|||||
AAATGG 190

76

Application US/10242535A
JS20040013663A1
ION:
iroGene Inc.

76

ION: Compositions and Methods Relating to Osteoarthritis
4231/2005
TION NUMBER: US/10/242,535A
DATE: 2002-09-12
ON NUMBER: US 10/085,783
TE: 2002-02-28
ON NUMBER: US 60/305,340
TE: 2001-07-13
ON NUMBER: US 60/275,017
TE: 2001-03-12
ON NUMBER: US 60/271,955
TE: 2001-02-28
ON NOS: 58994
In version 3.2

76

feature
... (213)
ION: n is a, c, g, or t
76

13.7%; Score 188.4; DB 15; Length 213;
urity 99.5%; Pred. No. 1.le-41;
nservative 0; Mismatches 1; Indels 0; Gaps 0;

ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 1243
ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 60
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 1303
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 120
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTGTGACAAATGT 1363
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTATTATTGTGACAAATGT 180

AAATGG 1373
|||||
AAATGG 190

plication US/09983965
0137160A1
ON:
n, Wesley C.
Nengbing
t, John C.
ialagan, Nagappan
ON: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
ON: MUSCLE AND FAT DEPOSITION
37-21 (10297)C
ION NUMBER: US/09/983,965

106:25:17 2004

us-09-245-198a-3.rnpb

|||||
ATTCCTGTGGATTTTAAACAGATATTATTTATTATTATTGTGACAAATGT 180
|||||
AAATGG 1373
|||||
AAATGG 190

76

Application US/10242535A
JS20040013663A1
ION:
iroGene Inc.

76

ION: Compositions and Methods Relating to Osteoarthritis
4231/2005
TION NUMBER: US/10/242,535A
DATE: 2002-09-12
ON NUMBER: US 10/085,783
TE: 2002-02-28
ON NUMBER: US 60/305,340
TE: 2001-07-13
ON NUMBER: US 60/275,017
TE: 2001-03-12
ON NUMBER: US 60/271,955
TE: 2001-02-28
ON NOS: 58994
In version 3.2

76

feature
... (213)
ION: n is a, c, g, or t
76

13.7%; Score 188.4; DB 15; Length 213;
urity 99.5%; Pred. No. 1.le-41;
nservative 0; Mismatches 1; Indels 0; Gaps 0;

ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 1243
ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 60
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 1303
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 120
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTATTATTGTGACAAATGT 1363
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTATTATTATTGTGACAAATGT 180

AAATGG 1373
|||||
AAATGG 190

plication US/09983965
0137160A1
ON:
n, Wesley C.
Nengbing
t, John C.
ialagan, Nagappan
ON: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
ON: MUSCLE AND FAT DEPOSITION
37-21 (10297)C
ION NUMBER: US/09/983,965

106:25:17 2004

us-09-245-198a-3.rnpb

|||||
ATTCCTGTGGATTTTAAACAGATATTATTTATTATTATTGTGACAAATGT 180
|||||
AAATGG 1373
|||||
AAATGG 190

76

Application US/10242535A
JS20040013663A1
ION:
iroGene Inc.

76

ION: Compositions and Methods Relating to Osteoarthritis
4231/2005
TION NUMBER: US/10/242,535A
DATE: 2002-09-12
ON NUMBER: US 10/085,783
TE: 2002-02-28
ON NUMBER: US 60/305,340
TE: 2001-07-13
ON NUMBER: US 60/275,017
TE: 2001-03-12
ON NUMBER: US 60/271,955
TE: 2001-02-28
ON NOS: 58994
In version 3.2

76

feature
... (213)
ION: n is a, c, g, or t
76

13.7%; Score 188.4; DB 15; Length 213;
urity 99.5%; Pred. No. 1.le-41;
nservative 0; Mismatches 1; Indels 0; Gaps 0;

ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 1243
ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 60
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 1303
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 120
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTATTATTGTGACAAATGT 1363
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTATTATTATTGTGACAAATGT 180

AAATGG 1373
|||||
AAATGG 190

plication US/09983965
0137160A1
ON:
n, Wesley C.
Nengbing
t, John C.
ialagan, Nagappan
ON: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
ON: MUSCLE AND FAT DEPOSITION
37-21 (10297)C
ION NUMBER: US/09/983,965

106:25:17 2004

us-09-245-198a-3.rnpb

|||||
ATTCCTGTGGATTTTAAACAGATATTATTTATTATTATTGTGACAAATGT 180
|||||
AAATGG 1373
|||||
AAATGG 190

76

Application US/10242535A
JS20040013663A1
ION:
iroGene Inc.

76

ION: Compositions and Methods Relating to Osteoarthritis
4231/2005
TION NUMBER: US/10/242,535A
DATE: 2002-09-12
ON NUMBER: US 10/085,783
TE: 2002-02-28
ON NUMBER: US 60/305,340
TE: 2001-07-13
ON NUMBER: US 60/275,017
TE: 2001-03-12
ON NUMBER: US 60/271,955
TE: 2001-02-28
ON NOS: 58994
In version 3.2

76

feature
... (213)
ION: n is a, c, g, or t
76

13.7%; Score 188.4; DB 15; Length 213;
urity 99.5%; Pred. No. 1.le-41;
nservative 0; Mismatches 1; Indels 0; Gaps 0;

ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 1243
ACCCCACTCAGGACTAAGAGGGGCTGGACCTGGCGGAGGAGCCAAAGAGAC 60
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 1303
CTAGGCCAGGAGTTCCTCCAAATGTGAGGGGCGAGAAACAGACAGCTCCTCCCTT 120
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTATTATTGTGACAAATGT 1363
TTCCCTGTGGATTTTAAACAGATATTATTTATTATTATTATTATTATTGTGACAAATGT 180

AAATGG 1373
|||||
AAATGG 190

plication US/09983965
0137160A1
ON:
n, Wesley C.
Nengbing
t, John C.
ialagan, Nagappan
ON: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
ON: MUSCLE AND FAT DEPOSITION
37-21 (10297)C
ION NUMBER: US/09/983,965


```

QY 25 GAVRQAQPPAPMAARRSRR-----RRRGPGTALLVPLALGLGLALA
Db 19 GAIKOKS-----MAVEKNRRALGIDGNVTVVGVGKALPQVSRPITRGF-----
QY 74 -----LAVVSLGSRASLS-----AQEPAQBELVAEBEDODPSLNPO
Db 70 ANAEAAAAENNNKNSLAVNAGADGALPIKKNVAVVPQKKTVKSPQETIIETSPDT
QY 115 DPAPFLNLRPRKS-----APGKRKTRARRATAAHYEVHPRPGQ-----DGAQAQY
Db 130 --APVLEKEITGERSLKKKAPTJLTSTLTARSKAASV-VTKPKQEIVIDDAVDNN
QY 166 SWEF-----ARINSSPLRY--NROIGFIVTRAGLYLYCQVHFD--EGF
Db 187 VEYVEDMYKFKYSAENDSRPHDYMDSQPEINEKM--RAILIDLWLVQVHKFELSP
QY 213 KLDLLVDGVIALRC-----LSEFSATASSIGPOLRLCQVSGLLALRPGS 257
Db 245 TIN-IVRYLASKTTSRRLEQLIGMSSMLIASKYBEINAPEVNDLVCSIDGS 295

RESULT 8
T36946
probable cation-transporting ATPase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug
C:Accession: T36946
R:Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: 221607
A:Accession: T36946
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-776 <SEE>
A:Cross-references: EMBL:AL109862; PIDN: CAB53131.1; GSPDB: GN000070; SCO
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCORDB: SCJ1.13
C:Superfamily: ATPase nucleotide-binding domain homology
F:442-585/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 6.5%; Score 93.5; DB 2; Length 776;
Best Local Similarity 24.1%; Prd; No. 12;
Matches 63; Conservative 30; Mismatches 103; Indels 65; G; G;

QY 47 RRGEPTGTVLAVPLALGLGLALA--CLGLLLA-VVSLGSRASLSAQEPAQBELVAEE
Db 73 RRGHAGVDLIVLALUGGTLVAGEYIAGVLIALMTAGTGTLEGAQRASHDHLHALL
QY 104 SELNPQTESQDPAPPLNLRPRRSAPKGRKTRARRATAAHYEVHPRPGDGAQA
Db 133 RSARRRTGDG-----VVR-----VPLSEITAGDALVVGPEVVP-----
QY 164 TVSGWE---EARINSSPLRYNRQIGE----FIVTRAGLYLYCQVHFDGKAVYLI
Db 170 RVESTEAVLDESVLGEIQLVTRQREGARGAVNAGAFDL-----RATAII
QY 217 LVQGVIALRCLEFSATASSIGPOLRLCQ-----VSGLLALRPGSSLR
Db 221 TYAGIVRL-----AQQAGASAPVVRADRYAAWFLPLATAALAWLVSGSAVRF
QY 261 -IRTLPAWHLKAAPFLTYFGL 280
Db 275 LVVATPCPILLAAPVAVVSGL 295

RESULT 9
T115838
hypothetical protein C54D2.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-
C:Accession: T115838
A:Minx, P.

```

EMBL Data Library, October 1995
 e sequence of *C. elegans* cosmid C54D2.
 r: Z18415

38
 nary; translated from GB/EMBL/DBJ
 DNA
 7 <MIN>

9: EMBL:U37548; NID: g1017804; PID: g1017809; PIDN: AAA79201.1; CESP: C54D2
 2.5
 67/3; 86/3; 121/2; 199/3; 230/2; 308/2; 334/2; 370/2; 439/3; 470/3; 513
 dium channel protein

6.5%; Score 93.5; DB 2; Length 1657;

larity 23.2%; Pred. No. 29;

Conservative 38; Mismatches 99; Indels 71; Gaps 12;

LLVPLALGLALACLG-----LLAVVSLGSRAS-----LSAQEPAQEELVAEED 100

QTNPWALYFVALMTFGNYVLFNLVAILVEGFOESKEEKRLQLEEDARKQAVEED 940

SELPQTESQDPAPFLNLRVP-----RRSAPKGRKTRARRAIAAHYEVHPRPGQ 154

RELELIIAKTSPA--FNNGVAPAECTCQRPSP--ESPPRLLSANY--HSPER 994

QAGVDGTVSGWEARINSSPL-----RYNRQIGFIVTRAGLYLYCQV 202

-ANLDAIID--KRLVLRNSAPFDRSPVSEGRDSDRLNRHASLVLPVANGVYRQV 1051

-----FDEGKAVYKLDLVDGVALRCLE---ESATAASLGLPQLRLCQV 247

KASQELKQALAEKNEAKNTFVRKLLKTKLHNTEFS----- 1095

LALRPGSSLRIRTPWAHLKAAPFLTYF 278

FLMGPKNPLRIKCLQTTQKKWEDYVLF 1124

in Rv0497 - Mycobacterium tuberculosis (strain H37Rv)

terium tuberculosis

98 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

sch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

tes, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

14, 1998

R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

ng the biology of Mycobacterium tuberculosis from the complete genome

: A70500; MUID: 98295987; PMID: 9634230

15

lary; nucleic acid sequence not shown; translation not shown

INA

<COL>

: GB: Z77162; GB: AL123456; NID: g3261606; PIDN: CAB00923.1; PID: e255036;

urce: strain H37Rv

6.4%; Score 93; DB 2; Length 310;

arity 24.4%; Pred. No. 4.8;

onservative 25; Mismatches 103; Indels 120; Gaps 18;

SLGSRDGA---VRQAQPPAPMAARRSQRRG-----R 47

SSGNRQISVAELLARQGTGAP--ARRRRRRGSDAITVAELTGETPIIRDHHH 63

GTALLVPLALGLALACGLLLAVVSLGSRASLSAQEPAQEELVAEE----- 99

AHASQSPAANGR-----VQGEAAPQSPAEPPVAEQ--VAEPTFTVWS 109

SELPQTESQDPAPFLNLRVP-----SAPKGRKTRARRAI---AAHY----- 146

Db 110 QPERWPKSPQDRRESGPELSEYPRLRTHSDRAPAGPPSGAEHMSPPDVEH;
 QY 147 -----EVHPRPGDG-----AQAGVDGTVSGWEARINSSS---
 Db 170 DVLDTVEGEAEAEVREAOQGRGERHAAAGAAAGTVEGDGAARVARALDVI
 QY 183 RQIGERTVTR-----AGLYLYCQVHFE-----GKAVYKLDL-----LVDC
 Db 230 ---GALVLIQSILAVAFGAGLF-----IAFDQLWRMNSIVALVLSVMVILGLVVS
 QY 226 CLEFSAT-----AASSLGPQLRLCQ 246
 Db 282 KTDIASTLIAVAVGALITLGP-LALLQ 308

RESULT 11

B34768

ORF5 protein - Orf virus (strain NZ2)

C;Species: Orf virus

C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 08-Oct

C;Accession: B34768

R;Fraser, K.M.; Hill, D.F.; Mercer, A.A.; Robinson, A.J.

Virolology 176, 379-389, 1990

A;Title: Sequence analysis of the inverted terminal repetition in the

A;Reference number: A34768; MUID: 90266454; PMID: 2129563

A;Accession: B34768

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-351 <FRA>

A;Cross-references: GB:M30023; EMBL:M37623; NID: g332561; PIDN: AAA4678

Query Match

Best Local Similarity 6.4%; Score 92; DB 2; Length 351;

Matches 45; Conservative 18; Mismatches 73; Indels 60;

QY 14 PLPRSLGSRDGGAVRAQAPPAARRSQ--RRGRGEGTALLVPLALGLAL

Db 211 PLPRRAAR---GQRGQPPPRARRAQPRRRAPRAAG-----

QY 73 LLAVVSLGSRASLSAQEPAQEELVAEEDQPSLNPQTESQDPAPFLNLRVPR

Db 248 -----ARRGRGPAPROQQRPRVORAAAAAQRRAQR

QY 133 GKTRARRAIAHYB-VHPRGQDGAQAGVDGTVSGWEARINSSPLRYNRQIG

Db 284 PRVVRARRARRQORAHQR--RRGRARRTRCSTS-----RVVSKD-----SREVG

QY 192 RAGLYLYCQVHFEFG 207

Db 333 KERRYIRRVLLHFEFG 348

RESULT 12

T35203

probable two-component sensor - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 31-Jan

C;Accession: T35203

R;Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.J

submitted to the EMBL Data Library, April 1998

A;Reference number: Z21571

A;Accession: T35203

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-566 <SEE>

A;Cross-references: EMBL:AL022374; PIDN: CAA18527.1; GSPDB: GN00070; SCC

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SC5B8.19C

C;Superfamily: two-component sensor histidine kinase; sensor histidine

Query Match 6.4%; Score 92; DB 2; Length 566;

arity 24.5%; Pred. No. 11;
 onservative 24; Mismatches 99; Indels 80; Gaps 14;
 MAARRSQR-----RGRGEP-----GTALLVPLALGLGALACGLLAVVSLGS-- 81
 RTVAMGSTPPVRLRLGLPRVFSQVLLMQLAAGVAVLATGLFLA--PLGDQL 59
 --RASLSAQEPAQBELVAEBDDQPSLNPQTESQDPAPFLNLRVRRSAPKGRK 135
 MRRALATAQTAAQPPQVVRD-----LRTTRTPANGPVQRE 98
 RRATAAHYEV-----HPRPCDQAQAGVD--GTVSGWEEARINSSS----- 177
 REATRAEYVVVMDRQGVNWSHTDPERIGEVVSTDPQALAGREVMEIDDTLGRSA 158
 PLRYNRQIGEFI-VTRAGLYLYCO--VHFDEGKAVY-----LKLDDLVDGVLALR 225
 PLRDGD--GEIVGAVSVGIADSVEARLIHALPGLFAYAGGALAVGALASWIIISR 216
 IFSATAASSLGPQLRLCQVSGLLALR 254
 YTFDLAFS-----DIAGLLAER 236

ain Rv1219c - Mycobacterium tuberculosis (strain H37RV)
 aterium tuberculosis
 98 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 11
 sch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 44, 1998
 R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 ing the biology of Mycobacterium tuberculosis from the complete genome
 r: A70500; MUID:98295987; PMID:9634230
 11
 nary; nucleic acid sequence not shown; translation not shown
 DNA
 <COL>
 s: GB:Z93777; GB:AL123456; NID:G3261726; PIDN:CAB07841.1; PID:ei299832;
 urce: strain H37RV

6.3%; Score 91.5; DB 2; Length 212;
 larity 24.5%; Pred. No. 4.1;
 Conservative 26; Mismatches 78; Indels 59; Gaps 9;
 (GEPGTALLVPLALGLG-----ALACGLLLAVVSLGSRASLSAQEPAQBELVAEB-- 99
 IG-----FGVGLRAIAEAGVSAALVIHFHFSKEGL---RKACDDFVAEEIR 66
 -----DQDPSLNPQTESQDPAPFLNLRVRRSAPKGRKTRARRAJAAHYEVHPR 151
 KAAALKSNDPTTWLAQMAETSYAPLMAYLVRSMSQSGELAKMLWQKI----- 117
 QDGAQAGVDGTVSGWEEARINSSSPLRYNRQIGEFI-VTRAGLYLYCQVHFD----- 205
 -DNAEEYLD-----EGVRAGTVKPSRDPARARFLAITGGGFFLLYLMHENPTDLR 168
 -----EGKAVYLKDLLVDGVLALRCL--EFSATA 234
 LRQYAHDMVLPSELYVTEGLLADRAMVEAFLAE 204

tein PA3305 [imported] - Pseudomonas aeruginosa (strain PA01)
 monas aeruginosa
 :000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 .231

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an c
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: D83231
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-664 <STO>
 A;Cross-references: GB:AE004753; GB:AE004091; NID:G9949433; PIDN:AAG06
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: PA3305

Query Match 6.3%; Score 91.5; DB 2; Length 664;
 Best Local Similarity 25.8%; Pred. No. 15;
 Matches 77; Conservative 28; Mismatches 117; Indels 77;
 QY 28 RQAQPPAPMAARR-----SOR-----RGRGEPGTALLVPLALGLGALACI
 Db 194 RQRPQGLLNALSKIVEVDQRDHAWFEGERRRAGALALLSRDLLSL-
 QY 75 AVVSLGSRASLSAQEPAQEE-----LVAEEDQDPSELNPQTESQDPAPFLNRL
 Db 246 ARGVARQARLSEEEERRVERVWLAALASALEGTDPASMQALRELAQVA-
 QY 129 SAPGKRKTRARRAIAAHYEVHPRPDQDGAQAGVDGTVSGWEEARINSSSPLRYNR
 Db 300 SNDQ-RYLLTRCSVLLKAVN-----AEKGMRAVASGEVGRVGSAGTLSWHR
 QY 187 EFIVTRAGL-----YLYL-----CQVHFDGKAVYLKDLILV
 Db 352 LFYGTTSALALLGLSVYVIYITAWPAASGAMLLAAVVCSEFANRDNVAIGLSFLR
 QY 224 LRCLFEFSATAASSLGPQLRLCQVSG--LIALRPGSSLRIRTPMAHLKAPFLI
 Db 412 I-----PAAMLVISQWLLPQWNGFPLLCIAMGVPLFPATILGMVAVPTAGTAI

RESULT 15
 H83044
 2,4-dienoyl-CoA reductase FadH2 PA4814 [imported] - Pseudomonas aeru
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-D
 C;Accession: H83044
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, J.
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: H83044
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-661 <STO>
 A;Cross-references: GB:AE004894; GB:AE004091; NID:G9951076; PIDN:AAG
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: fadH2; PA4814
 C;Superfamily: Methylophilus methylotrophus W3A1 trimethylamine dehy

Query Match 6.3%; Score 91; DB 2; Length 681;
 Best Local Similarity 26.3%; Pred. No. 17;
 Matches 79; Conservative 31; Mismatches 86; Indels 104;
 QY 6 FEISARRPLPRSLGRDGGAVRQAQPPAP-----MAARRSQRRGR-----
 Db 442 FRVRLERLGVLDLGRH-----VRQELDQDFDDVVVATGICQPRRPRIDGIGPT
 QY 48 --RGEP--GTALLVPLALGLGALACGLLLAVVSLGSRASLSAQEPAQBELVAE
 Db 498 VLRGAPVGARVAIVGAGGIGFDVA--AFLVAAPSDG-----QPRALGEWLAI

6:25:24 2004

us-09-245-198a-4.rspt

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

1 search, using sw model

il 7, 2004, 17:41:27 ; Search time 39.6149 Seconds
(without alignments)
2261.954 Million cell updates/sec

09-245-198A-4
4
SLLDPEISARRLPRLSLG.....PWAHLKAAPFLTYFGLFQVH 284

SUM62

op 10.0 , Gapext 0.5

7041 seqs, 315518202 residues

s satisfying chosen parameters: 1017041

th: 0

th: 2000000000

nimum Match 0%

ximum Match 100%

sting first 45 summaries

TREMBL_25:*

sp archaea:*

sp bacteria:*

sp fungi:*

sp human:*

sp invertebrate:*

sp mammal:*

sp mhc:*

sp organelle:*

sp phage:*

sp plant:*

sp rodent:*

sp virus:*

sp vertebrate:*

sp unclassified:*

sp virus:*

sp bacteriaph:*

sp archaea:*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

xy	ch	Length	DB	ID	Description
9	410	11	Q8BXS2	Q8bxs2 mus musculu	
4	330	4	Q81ZK7	Q81zk7 homo sapien	
0	261	5	Q8MRW2	Q8mrw2 drosophila	
0	325	5	Q9V5G2	Q9v5g2 drosophila	
0	409	5	Q8MY88	Q8my88 drosophila	
0	415	5	Q8MUJ1	Q8muj1 drosophila	
8	409	5	Q8IGD3	Q8igd3 drosophila	
2	398	6	Q8MK49	Q8mk49 sorex ciner	
2	532	16	Q8ZHP6	Q8zhp6 streptomyc	
1	565	16	Q8RY66	Q8ry66 streptomyc	
0	955	10	Q84T85	Q84t85 oryza sativ	
0	967	10	Q7XLL4	Q7xll4 oryza sativ	
0	643	16	Q9KZ17	Q9kz17 streptomyc	
0	375	16	Q9RRH5	Q9rrh5 deinococcus	
0	893	16	Q81ZX0	Q81zx0 streptomyc	
9	611	11	Q8KOM8	Q8kom8 mus musculu	

17	100	6.9	850	11	Q9JJ15	Q9jj15 mus
18	100	6.9	850	11	Q8OXI6	Q8oxi6 mus
19	99.5	6.9	378	16	Q9AAB9	Q9aab9 caul
20	99	6.9	330	4	Q8N5L1	Q8n5l1 homo
21	99	6.9	614	4	Q7Z4K2	Q7z4k2 homo
22	99	6.9	915	4	Q7Z5I1	Q7z5i1 homo
23	98	6.8	206	16	Q9S2W5	Q9s2w5 stre
24	97.5	6.8	694	16	Q82FL1	Q82fl1 stre
25	97.5	6.8	1560	4	Q96JP2	Q96jp2 homo
26	97	6.7	408	10	Q8S5I5	Q8s5i5 oryz
27	97	6.7	926	4	Q9NYA0	Q9nya0 homo
28	97	6.7	1058	4	Q9Y4G2	Q9y4g2 homo
29	96.5	6.7	975	11	Q8BWB1	Q8bwb1 mus
30	96	6.6	629	10	Q8SIA6	Q8sia6 oryz
31	95.5	6.6	536	4	Q9HB96	Q9hb96 homo
32	95.5	6.6	655	16	Q9FBR7	Q9fbr7 stre
33	95	6.6	340	16	Q7WFL3	Q7wfl3 bord
34	95	6.6	390	2	Q8KW28	Q8kw28 ruege
35	95	6.6	810	16	Q82K60	Q82k60 stre
36	94.5	6.5	748	5	Q8T2Y0	Q8t2y0 trypa
37	94.5	6.5	1696	11	Q9WTR8	Q9wtr8 ratt
38	94	6.5	340	16	Q7W3N7	Q7w3n7 bord
39	94	6.5	340	16	Q7W0H0	Q7w0h0 bord
40	94	6.5	448	16	Q886A1	Q886a1 pseu
41	94	6.5	937	16	Q93JDI	Q93jdi stre
42	94	6.5	1910	10	Q7XU19	Q7xu19 oryz
43	93.5	6.5	776	16	Q9RJ01	Q9rj01 stre
44	93.5	6.5	854	16	Q9F2P0	Q9f2p0 stre
45	93.5	6.5	1038	10	Q9AS09	Q9as09 oryz

ALIGNMENTS

RESULT 1

Q8BXS2	ID	Q8BXS2	PRELIMINARY;	PRT;	410 AA.
AC	Q8BXS2;				
DT	01-MAR-2003 (Tremblrel. 23, Created)				
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)				
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)				
DE	Tumor necrosis factor.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Retina;				
RX	MEDLINE=22354683; PubMed=12466851;				
RA	The FANTOM Consortium,				
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;				
RT	"Analysis of the mouse transcriptome based on functional annotation				
RT	60,770 full-length cDNAs."				
RL	Nature 420:563-573 (2002);				
DR	EMBL; AK044387; BAC31897.1;				
DR	PIR; PT0714; PT0714.				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.				
DR	GO; GO:0008955; P:immune response; IEA.				
DR	InterPro; IPR006052; TNF_family.				
DR	InterPro; IPR008983; TNF_like.				
DR	SMART; SM00207; TNF; 2.				
DR	PROSITE; PS00251; TNF_1; 1.				
DR	PROSITE; PS50049; TNF_2; 2.				
SQ	SEQUENCE 410 AA; 45881 MW; 590A4B74C33FB8D4 CRC64;				

Query Match 64.9%; Score 937.5; DB 11; Length 410;
Best Local Similarity 79.1%; Pred. No. 1.3e-73;
Matches 193; Conservative 11; Mismatches 29; Indels 11; G;

QY

36 MAARRQRGRGRGEGTALLVPLALGLALACLLGLLVAVSLGSRASLSA-QEP;
|||||

```

RSQRRRGRRGPTALLVPLVSLGLALACLLLVVSLGSWATLSAQEPFSQBE 60
: : : : : : : : : : : : : : : : : : : : : : : : : :
:EDDPSLNPTQTESQDPAPFLNRLVRRPAPKGRKTRARRAIAAHVEVHPRPQ 154
: : : : : : : : : : : : : : : : : : : : : : : : : :
:DRPEPELNPTQTESQDVVPFLEQLVRRPAPKGRKARPRRAIAAHVEVHPRPQ 120
: : : : : : : : : : : : : : : : : : : : : : : : : :
:AGVDGTGSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAVYLK 214
: : : : : : : : : : : : : : : : : : : : : : : : : :
:AGVDGTGSGWEETKINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAVYLK 180
: : : : : : : : : : : : : : : : : : : : : : : : : :
:JGVLAIRCLESFASATASGLGQLRLCQVSGLLALR-----PGSLRLIRTL 265
: : : : : : : : : : : : : : : : : : : : : : : : : :
:NGVLAIRCLESFASATASGLGQLRLCQTE-LQSLRREVSLRQSGPQKQGRP 239
: : : : : : : : : : : : : : : : : : : : : : : : : :
: 269
: 243

PRELIMINARY; PRT; 330 AA.

(TREMBLrel. 23, Created)
(TREMBLrel. 23, Last sequence update)
(TREMBLrel. 25, Last annotation update)

(Human)
Mammalia; Chordata; Craniata; Vertebrata; Eutelestomi;
Mammalia; Primates; Catarrhini; Hominoidea; Homo.
306;

4 N.A.
9924; PubMed=12411489;
B.; Medema J.P., Lopez-Fraga M., Lozano J.C.,
M., Picard A., Martinez-A C., Garcia-Sanz J.A.,
is hybrid mRNA encodes TWE-PRIL, a functional cell surface
fusion protein.";
11-5720(2002);
1; AAL90443.1;
2; C:membrane; IEA.
34; F: tumor necrosis factor receptor binding; IEA.
35; P: immune response; IEA.
3006052; TNF family.
3008983; TNF_like.
3; TNF; 1.
37; TNF; 1.
3251; TNF; 1.
3049; TNF; 2.
30 AA; 36588 MW; PC6F3BCA29C029AE CRC64;

Larity 58.4%; Score 844; DB 4; Length 330;
Conservative 100.0%; Pred. No. 1.5e-65; Indels 0; Gaps 0;

RSQRRRGRRGPTALLVPLVSLGLALACLLLVVSLGSWATLSAQEPQEL 95
: : : : : : : : : : : : : : : : : : : : : : : : : :
RSQRRRGRRGPTALLVPLVSLGLALACLLLVVSLGSWATLSAQEPQEL 60
: : : : : : : : : : : : : : : : : : : : : : : : : :
EDDPSLNPTQTESQDPAPFLNRLVRRPAPKGRKTRARRAIAAHVEVHPRPQ 155
EDDPSLNPTQTESQDPAPFLNRLVRRPAPKGRKTRARRAIAAHVEVHPRPQ 120
: : : : : : : : : : : : : : : : : : : : : : : : : :
AGVDGTGSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQ 201
: : : : : : : : : : : : : : : : : : : : : : : : : :
AGVDGTGSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQ 166

PRELIMINARY; PRT; 261 AA.

```

```

DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE SD182860.
DE BIGER OR CG12919.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY119233; AAM51093.1;
DR FlyBase; FBgn0033483; eiger.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR006052; TNF family.
DR InterPro; IPR008983; TNF_like.
DR SMART; SM00207; TNF; 1.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS50049; TNF_2; 1.
DR PROSITE; PS50049; TNF_2; 1.
SQ SEQUENCE 261 AA; 29780 MW; 13B6D5A04EC9122C CRC64;

Query Match 8.0%; Score 116; DB 5; Length 261;
Best Local Similarity 25.4%; Pred. No. 0.048;
Matches 57; Conservative 32; Mismatches 97; Indels 38;

QY 82 RASLSAQEPQAEELVAE-----EDQPSLNPTQTESQDPAPFLNR-----
DB 54 RKSRSIADVRNEEQNIQGNHTELOEKSNEATSKES--PAPLHRRRMHSRHHL
QY 128 RSAPKGRKTRARRAIAAHVEVHPRPQDCAQAGVDGTGSGWEARINSSPLRYN
DB 112 ESSLARSSEDSRP--AAHFLLSSRRRHQSGM-GYHGDVYIGNDNERNYSQG-HFQ
QY 188 FIVTRAGLYLYCQV-----HFDEGKAVYLKDLLVDGLARCLESFASATASS
DB 168 LTVTNTGLVYVYAQICYNNSHDQGFIVE-----QGDTPLQCLN---TVPTN
QY 243 RLCQVSGLLALRPGSSLRITL---PWAHLKAAPFLTYGFLFQV 283
DB 218 HTCHTSGLIHLERNRIHLKDIHNDRNAVLREGNRSYFGIFKV 261

RESULT 4
Q9V5G2
ID Q9V5G2 PRELIMINARY; PRT; 325 AA.
AC Q9V5G2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CG12919 protein.
GN BIGER OR CG12919.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.]

```



```

923; PubMed=12176339;
n M., Basler K.;
TNF Signaling Mechanisms. JNK-Dependent Apoptosis
Eiger, the Drosophila Homolog of the TNF Superfamily.";
2:1263-1268(2002).
N.A.
938; PubMed=12894227;
Maaty W.S., Chen P., Tomar R.S., Eby M.T., Chapo J.,
ore N., Zachariah S., Sinha S.K., Abrams J.M.,
s receptor, Wengen, comprise a TNF-like system in
860-4867(2003).
6; AAM76710.1; -.
1; AAM66763.1; -.
0033483; eiger.
0; C:membrane; IEA.
5; P:tumor necrosis factor receptor binding; IEA.
4; P:immune response; IEA.
006052; TNF family.
008983; TNF-like.
7; TNF; 1.
251; TNF 1; 1.
049; TNF 2; 1.
5 AA; 46918 MW; E087A26DE222D2BF CRC64;
8.0%; Score 116; DB 5; Length 415;
arity 25.4%; Pred.No.0.086; 97; Indels 38; Gaps 10;
onservative 32; Mismatches
SAQPAQELVAE---EDQPSLNPQTERSQDPAPFLNR-----LVRPR 127
SIADVRNEEQNIQNHTELOEKSNEATSKES--PAPLHRRMHSRRHLLVRKG 265
KGRTRARRATAAHYEVHPRPGQDGAQAGVDGTSGWEEARINSSPLRYNRQIGE 187
SARSEDSP--AAHFLSSRRHQGSM-GYHGDVYIGNDNERNYSQG-HFQTRDGV 321
RAGVLYLYCOV-----HFDEKAVYKLDLLVDGVLALRCLEEFSAATASLGPQL 242
NTGYYVYQAQICYNNSHDQNGFIVF-----QGDTPFLQCLN----TPTNMPHKV 371
VSGLLALRPGSSSLRIRTL---PWAHLKAAPELTYFGLFQV 283
TSGLIHLENERIHLKDIHNDNAVLREGNRSYFGIFKV 415
RELIMINARY; PRT; 409 AA.
TREMBLrel. 23, Created)
TREMBLrel. 23, Last sequence update)
TREMBLrel. 25, Last annotation update)
lanogaster (Fruit fly).
tazaa; Arthropoda; Hexapoda; Insecta; Pterygota;
lopterygota; Diptera; Brachycera; Muscomorpha;
Drosophilidae; Drosophila.
1 N.A.
Brokstein P., Hong L., Agbayani A., Carlson J.,
lavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
mazalez M., Guarin H., Kronmiller B., Li P., Liao G.,
ungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
uanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
V-2002) to the EMBL/GenBank/DBJ databases.
18; AAN71595.1; -.
DR FlyBase; FBgn0064801; BcdNA:RHS1659.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR006052; TNF family.
DR InterPro; IPR008983; TNF-like.
DR SMART; SM00207; TNF; 1.
DR PROSITE; PS00251; TNF 1; 1.
DR PROSITE; PS50049; TNF 2; 1.
DR PROSITE; PS50049; TNF 2; 1.
SQ SEQUENCE 409 AA; 46401 MW; FC2E9BD9E012D257 CRC64;
Query Match 7.8%; Score 113; DB 5; Length 409;
Best Local Similarity 23.7%; Pred.No.0.15;
Matches 54; Conservative 36; Mismatches 94; Indels 44; G
QY 84 SLSAQEPAQEL-----VAEEDQPSLNPQTEE-----SQDPAPFLNLVR
Db 198 SYNAAKKKKKRSKRSIADVRNEEQNIQNHTELOEKSNEATSKERAPLHRR---
QY 130 APKGRKTRARRA-----IAAHYEVHPRPGQDGAQAGVDGTSGWEEARINSSPL
Db 254 HSRHLLVRKARSEDSPAAHFLSSRRHQGSM-GYHGDVYIGNDNERNYSQG-
QY 184 QIGEFIVTRAGLYLYCOV-----HFDEKAVYKLDLLVDGVLALRCLEEFSAATP
Db 312 RDGVLTVTNTGLYVYVYQAQICYNNSHDQNGFIVF-----QGDTPFLQCLN---TV
QY 239 GPQLRLQVSGLLALRPGSSSLRIRTL---PWAHLKAAPELTYFGLFQV 283
Db 362 PKVHTCHTSGLIHLENERIHLKDIHNDNAVLREGNRSYFGIFKV 409
RESULT 8
Q8MK49 PRELIMINARY; PRT; 398 AA.
AC Q8MK49;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha 2B adrenergic receptor (Fragment).
GN ADRA2B.
OS Sorex cinereus (Masked shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Soricidae; Soricinae; Sorex.
OX NCBI_TaxID=36803;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608557; PubMed=11743200;
RA Murphy W.J., Eizirik E., O'Brien S.J., Madsen O., Scally M.,
RA Douady C.J., Teeling E., Ryder O.A., Stanhope M.J., de Jong W.W.,
RA Springer M.S.;
RT "Resolution of the early placental mammal radiation using Bayesia
RT phylogenetics.";
RL Science 294:2348-2351(2001).
DR EMBL; AJ315936; CAC87000.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 398 398
SQ SEQUENCE 398 AA; 43576 MW; D57E67B689535E27 CRC64;
Query Match 7.2%; Score 104; DB 6; Length 398;
Best Local Similarity 25.2%; Pred.No.0.91;
Matches 55; Conservative 23; Mismatches 76; Indels 64; C

```

[illegible]

Qy	159	-----AGVDGTV-SGWPEARINSSPLRNYROI GFIVTRAGLYYL
		: : : : :
Db	428	TGIDPPVSYGTTVLACDGTGVTQWNSAYGNMA-----IVTAKOGTETWY
		: : : : :
Qy	204	FDEGKAVYKLIDLIVGVIALRCLEFSAATASSLGPQLRLCQVSGILLALRPGSLL
		: : : : :
Db	475	LSTKYASGITVRAGDPI-----AFSGNSGNSGTGPLH-----FEVRPAGS;
		: : : : :
Qy	264	LPW 266
Db	521	LPW 523
RESULT 10		
Q9KY366		PRELIMINARY; PRT; 565 AA.
ID	Q9KY366	
AC	Q9KY366;	
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	
DE	Putative peptidase	
DN	SC04798 OR SCD63A.09C.	
OS	Streptomycetes coelicolor.	
OC	Bacteria; Actinobacterii; Actinobacteridae; Actinomycetales;	
OC	Streptomycinae; Streptomycetaceae; Streptomycetes.	
NCBI TaxId=1902;		
[1]	SEQUENCE FROM N.A.	
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Brown S.P., Harris D.;	
RL	Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.	
[2]		
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RA	Cerdeno A.M.; Parkhill J., Barrell B.G., Rajandream M.A.;	
RL	Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.	
[3]		
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2);	
RX	MEDLINE=97000351; PubMed=8843436;	
RA	Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,	
RA	Kinashi H., Hopwood D.A.;	
KT	"A set of ordered cosmids and a detailed genetic and physical map	
RT	of the 8 Mb Streptomycetes coelicolor A3(2) chromosome.";	
RT	Mol. Microbiol. 21:77-96(1996).	
[4]		
RN	SEQUENCE FROM N.A.	
RC	STRAIN=A3(2) / M145;	
RX	MEDLINE=21996410; PubMed=12000953;	
RA	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,	
RA	Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,	
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M	
RA	Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.	
RA	Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.	
RA	Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,	
RA	Seeger K., Saunders D., Sharp S., Squares R., Taylor I.,	
RA	Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,	
RA	Hopwood D.A.;	
RT	"Complete genome sequence of the model actinomycete Streptomycetes	
RT	coelicolor A3(2).";	
RT	Nature 417:141-147(2002).	
RL	AL939121; CAB92661.1; "	
DR	GO: GO:0004222; P:metallopeptidase activity; IEA.	
DR	G0: GO:0006508; P:proteolysis and peptidolysis; IEA.	
DR	InterPro: IPR0012886; Peptide_M37.	
DR	Pfam: PF01551; Peptide_M37; I.	
SW	Complete proteome.	
Q9	SEQUENCE 565 AA; 58070 MW; 7D0418D480C6A284 CRC64;	

Query Match 7.1%; Score 103; DB 16; Length 565;
Best Local Similarity 21.2%; Pred. No. 1.7;
Matches 70; Conservative 33; Mismatches 99; Indels 128; G.

```

FEISA---RRLPLPSLGRDGGAVROAQPAPMAARRSQRRRRRGPGTALL-- 56
DQEBATPDARIPVARA-GSRAGARRRQP-----AKRS-----ALLTI 331
ALGLL---GLALACILGLLAVVSLGRSLASQAEPQAE----- 94
ACVMSVAGIAAASVSLTG--DEGTETAASAPDPGNAEAPVPKPSAANKLDTOLT 389
EED-----ODPELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHY 146
GADDFADRASRTQBRIDLKAEQDAEKRAAQEAARKERLPKLPVYKHGLSAYY 449
RPGDQG-----AOAGVDGTVSGWEARINSPLRVNRQICE 187
--GQAGINWMSHTGIDFPVLOGITVMAATDGTVR-----TQFN SAYCN 491
TA-GLYLYYQCVH-----FDEGKAVYKLDLLVDGVLALRCLFEFSATAAS 236
TAKDGTETWYCHLSSYQVPSGTTVKAGDAI-----AYS GDSGN 533
QRLRCQVSGLLALRPGSSLRIRTLPLW 266
HLH-----FEVRPAGGSDIDPLPW 556
PRELIMINARY; PRT; 955 AA.
TREMBLrel. 24, Created)
TREMBLrel. 24, Last sequence update)
TREMBLrel. 24, Last annotation update)
protein OSJNB0006008.10.
10.
(japonica cultivar-group).
ridiplantae; Streptophyta; Embryophyta; Tracheophyta;
; Magnoliophyta; Liliopsida; Poales; Poaceae;
; Oryzae; Oryza.
947;
[N.A.
pponbare;
uan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
as B.J.J., Sun B.B., Peterson J.J., Quackenbush J.,
zberg S.L., Fraser C.M.;
chromosome 3 BAC OSJNB0006008 genomic sequence."
Y-2002) to the EMBL/GenBank/DBJ databases.
[N.A.
pponbare;
R-2003) to the EMBL/GenBank/DBJ databases.
16; AAO66523.1; -.
15 AA; 105582 MW; E44E88C0FF71CC9C CRC64;
arity 7.0%; Score 101.5; DB 10; Length 955;
conservative 23; Mismatches 46; Indels 35; Gaps 6;
RRRLPLPSLGRDGGAVROAQPAPMAARRSQRRR--GRGEPGTALLVPLALGLG 64
QNRPLP-----GSLMRAPPPPPPTAEAPRQLPGAAASPATNTTAA----- 180
ACILGLLAVVSLGRSLASQAEPQAEEDDPSELNPOTESODPAPFL----- 120
3PVVILKGLVKPMQASIGRNPQSNE---DKDEDESE-----EEEEEGPVPDRA 232
-----NRLVRRRSAP 131

```

```

Db 233 TIEAIRAKRQLOQPRHAAP 252
RESULT 12
Q7XLL4
ID Q7XLL4 PRELIMINARY; PRT; 967 AA.
AC Q7XLL4;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE OSJNB00094P09.13 protein.
GN OSJNB00094P09.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Wu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen S.H.
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL731625; CA05074.1; -.
SQ SEQUENCE 967 AA; 106012 MW; 3FA9D0CCE245B970 CRC64;
Query Match 7.0%; Score 101.5; DB 10; Length 967;
Best Local Similarity 23.9%; Pred. No. 4.6;
Matches 70; Conservative 35; Mismatches 123; Indels 65;
QY 17 RSLGS-----RDGGAVRQ-----AQPPAPMAARRSQRRRRRGPGTALLVI
Db 352 RKLGGTTPPSPPRGGA VRASSRRRPEGAAPTSPQEGEKKKRLKRTGE-----
QY 63 LGLALACILGLLLA-----VVSLGRSLASQAEPQAEELVAEE---DOI
Db 400 ---TEPCRGNLISPPRWSFNPRSDVPSRHPKSCQSEAEPPAAAEERRRESI
QY 107 NPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAAHYVHPRPGDGAQAGVI
Db 457 ADRLREAEAEAREAR-VRQAEAEAEAREAEAREAEAEAEAREAEAREAEATA---
QY 167 CWEEARINSSPLRYNRQIGEFIVTRAGLYLYYQCVHFDGKAVYKLDLLVD---
Db 511 GAPWPRANTAVLDGFAQVEALRAEAELEAAWTRV---EEGRR---SVDAMVVEGI
QY 220 -GVIALRCLEFSATASSLPQLRCQVSGLLALRPGSSLRIRTLPLW-AHLK 2;
Db 566 RHVSELEARKAALAEIAREVEEERAAALISTAMVEAQDTLRLQHASWEAEIK 61
RESULT 13
Q9KZ17
ID Q9KZ17 PRELIMINARY; PRT; 643 AA.
AC Q9KZ17;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein SC02220.
GN SC02220 OR SC10B7.15.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.

```

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523</
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------

GenCore version 5.1.6
copyright (c) 1993 - 2004 Compugen Ltd.

a search, using sw model

il 7, 2004, 17:38:07 ; Search time 11.7171 seconds
(without alignments)
1262.081 Million cell updates/sec

09-245-198A-4

4
SLLDFFEISARLPLRLSLG.....PWAHLKAAPFLTYGLFQHV 284

SUM62

op 10.0 , Gapext 0.5

681 seqs, 52070155 residues

s satisfying chosen parameters: 141681

th: 0

th: 2000000000

nimum Match 0%

ximum Match 100%

sting first 45 summaries

issProt_42.*

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ch	Length	DB	ID	Description
1.8	249	1	TN12_HUMAN	O43508 homo sapien
1.6	225	1	TN12_MOUSE	O54907 mus musculus
1.6	272	1	TNF5_CHICK	O918d8 gallus gall
1.5	260	1	TNF5_CANFA	O97626 canis famil
1.4	254	1	TNF9_HUMAN	P41273 homo sapien
1.9	952	1	HDA7_HUMAN	O8wni4 homo sapien
1.7	441	1	CG22_ANTMA	P34801 antirrhinum
1.6	201	1	TNFB_MACEU	O9xt48 macropus eu
1.5	280	1	TNF6_MACMU	O9myl6 macaca mula
1.4	310	1	RHO_MICLU	P52154 micrococci
1.4	240	1	TN14_HUMAN	O11162 mycobacteri
1.4	280	1	TNFC_CERTO	O43557 homo sapien
1.4	902	1	NFC4_HUMAN	O9bdn1 cercocebus
1.3	760	1	MLH1_MOUSE	Q14934 homo sapien
1.3	814	1	CADF_HUMAN	O9jk91 mus musculus
1.2	707	1	JIPI_MOUSE	P55291 homo sapien
1.2	280	1	MDCB_KLEPN	O9wv19 mus musculus
1.2	316	1	TN11_MOUSE	P71422 klebsiella
1.1	278	1	TNFB_RAT	O35235 m tumor nec
1.1	281	1	TNF6_HUMAN	P36940 rattus norv
1.1	422	1	GF11_HUMAN	P48023 homo sapien
1.1	574	1	SEN3_HUMAN	O99684 homo sapien
1.0	197	1	TNFB_RABIT	O9h414 homo sapien
1.0	204	1	TNFB_BOVIN	P10154 cryptotagus
1.0	291	1	TN10_MOUSE	Q06600 bos taurus
1.0	250	1	TNFC_MACEU	P50592 mus musculus
1.0	139	1	YOFB_BACSU	O9xt47 macropus eu
1.0	205	1	TNFB_MARMO	P54467 bacillus eu
1.0	241	1	TN13_MOUSE	O9jmo9 marmota mon
1.0	777	1	METE_CAUCR	Q9d777 mus musculus
1.0	933	1	VGLB_HSVAL	Q9aaw1 caulobacter
1.0	928	1	VGLB_HSVBP	P17471 bovine herpesv

34	85.5	5.9	932	1	VGLB_HSVBC	P12640 bovir
35	85	5.9	372	1	LKXB_MOUSE	O88609 mus t
36	85	5.9	379	1	LKXB_HUMAN	O60663 homo
37	84.5	5.9	401	1	APOC_MYCTU	P95013 mycot
38	84	5.8	284	1	TLX2_HUMAN	O43763 homo
39	84	5.8	310	1	TNFC_MARMO	O9jmo10 marmc
40	84	5.8	310	1	MIS_FIG	P79295 sus e
41	84	5.8	703	1	ZM15_HUMAN	Q9h091 homo
42	84	5.8	825	1	ICP0_HSV2H	P28284 herpes
43	83.5	5.8	416	1	RAGE_BOVIN	Q28173 bos t
44	83.5	5.8	505	1	TUB_MOUSE	P50586 mus n
45	83.5	5.8	545	1	RTN2_HUMAN	O75298 homo

ALIGNMENTS

RESULT 1
TN12_HUMAN
ID TN12_HUMAN STANDARD; PRT; 249 AA.
AC O43508: O8WUZ7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 12 (TNF-related w
inducer of apoptosis) (TWEAK) (APO3 ligand).
GN TNFSF12 OR APO3L OR DR3LG.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND N-TERMINUS OF SOLUBLE FORM.
RC TISSUE=Fetal liver, and Tonsil;
RX MEDLINE=98070415; PubMed=9405449;
RA Chichepoteche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
Hession C., Garcia I., Browning J.I.;
RT "TWEAK, a new secreted ligand in the tumor necrosis factor family
weakly induces apoptosis."
RL J. Biol. Chem. 272:32401-32410(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal kidney;
RX MEDLINE=9828355; PubMed=9560343;
RA Marsters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,
Ashkenazi A.;
RT "Identification of a ligand for the death-domain-containing recept
Ap3."
RL Curr. Biol. 8:525-528(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.
Brownstein M.J., Uslidin T.B., Toshiyuki S., Carninci P., Prange C.
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H
Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley A.C., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

061: PubMed=10085077;
 ang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;
 s angiogenesis and proliferation of endothelial cells.";
 .. 274:8455-8459(1999).
 f apoptosis in some cell types. Mediates NF-KappaB
 n. May promote angiogenesis and the proliferation of
 al cells.
 Homotrimer (Potential).
 AR LOCATION: Type II membrane protein and secreted.
 ECIFICITY: Highly expressed in adult heart, pancreas,
 muscle, brain, colon, small intestine, lung, ovary,
 spleen, lymph node, appendix and peripheral blood
 es. Low expression in kidney, testis, liver, placenta,
 d bone marrow. Also detected in fetal kidney, liver,
 brain.
 soluble form derives from the membrane form
 lytic processing.
 Y: Belongs to the tumor necrosis factor family.
 Ref 3 sequence differs from that shown due to a
 t in position 125.

 OT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 dres a license agreement (See <http://www.isb-sib.ch/announce/>
 ail to license@isb-sib.ch).

 9; AAC51923.1; -;
 2; AAC39724.1; -;
 7; AAH19047.1; ALT_FRAME.
 1927; TNFSF12.

 7; C: integral to plasma membrane; TAS.
 2; P: receptor binding; TAS.
 7; P: induction of apoptosis; TAS.
 5; P: signal transduction; TAS.
 006052; TNF family.
 008983; TNF_like.
 7; TNF; 1.
 7; TNF; 1.
 251; TNF 1; FALSE_NEG.
 049; TNF2; 1.
 (ogenesis; Apoptosis; Transmembrane; Glycoprotein;
 1 249 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
 MEMBER 12, MEMBRANE FORM.
 94 249 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
 MEMBER 12, SECRETED FORM.
 1 21 CYTOPLASMIC (POTENTIAL).
 22 42 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 (POTENTIAL).
 43 249 EXTRACELLULAR (POTENTIAL).
 93 94 CLEAVAGE.
 39 139 N-LINKED (GLCNAC...).
 19 AA; 27216 MW; E660843361C28EBA CRC64;
 87.8%; Score 1268; DB 1; Length 249;
 arity 100.0%; Pred. No. 4e-92;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 RSQRGRGPGTALLVPLALGLALACUGLLAVSLGSRASLSAQEPAQEL 95
 |||||
 RSQRGRGPGTALLVPLALGLALACUGLLAVSLGSRASLSAQEPAQEL 60
 |||||
 EDDPSLNPQTESQDPAPFLNRLVPRRSAPGKTRARRAIAAHYVHPRPGD 155
 |||||
 EDDPSLNPQTESQDPAPFLNRLVPRRSAPGKTRARRAIAAHYVHPRPGD 120
 |||||

QY 156 GAQAGVDGTVSGWBEARINSSPLRYNRQIGFIVTRAGLYLYLYCQVHFDGKAVY
 DQ 121 GAQAGVDGTVSGWBEARINSSPLRYNRQIGFIVTRAGLYLYLYCQVHFDGKAVY
 QY 216 LLDVGVLAIRCLREFSATAASSLGPQLRLCOVSGLLALRPGSSLRIRTLPAWHLKA
 DQ 181 LLDVGVLAIRCLREFSATAASSLGPQLRLCOVSGLLALRPGSSLRIRTLPAWHLKA
 QY 276 TYFGLFQVH 284
 DQ 241 TYFGLFQVH 249

 RESULT 2
 TN12 MOUSE
 ID TN12 MOUSE STANDARD; PRT; 225 AA.
 AC 054907; 09CTP2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Tumor necrosis factor ligand superfamily member 12 (TNF-related
 DE inducer of apoptosis) (TWEAK) (Fragment).
 GN TNFSF12.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Peritoneal macrophage;
 RX MEDLINE=98070415; PubMed=9405449;
 RA Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
 RA Hession C., Garcia I., Browning J.L.;
 RT "TWEAK, a new secreted ligand in the tumor necrosis factor family;
 RT weakly induces apoptosis".
 RL J. Biol. Chem. 272:32401-32410(1997).
 RN [2]
 RP SEQUENCE OF 93-225 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washit
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.I.
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmi
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -!- FUNCTION: Binds to FN14 and possibly also to TNFSF12/AP03. I
 CC inducer of apoptosis in some cell types. Promotes angiogenes
 CC the proliferation of endothelial cells. Mediates NF-KappaB
 CC activation (By similarity).
 CC -!- SUBUNIT: Homotrimer (Potential).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted
 CC similarity).
 CC -!- TISSUE SPECIFICITY: Widely expressed.
 CC -!- PTM: The soluble form is produced from the membrane form by
 CC proteolytic processing (By similarity).
 CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a col.

STANDARD; PRT; 260 AA.

Rel. 40, Created

Rel. 40, Last sequence update)

Rel. 41, Last annotation update)

s factor ligand superfamily member 5 (CD40 ligand).

OLG OR CD40L.

ris (Dog).

tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
heria; Carnivora; Fissipedia; Canidae; Canis.

15;

N.A.

illett B.J.;
properties of canine CD40L";
G-1998) to the EMBL/GenBank/DBJ databases.
Cytokine that binds to TNFRSF5. Mediates B-cell
tion in the absence of co-stimulus as well as IgE
n in the presence of IL-4. Involved in immunoglobulin
tching (By similarity).

Homotrimer (By similarity).

AR LOCATION: Type II membrane protein. Also exists as an
ular soluble form (By similarity).

soluble form derives from the membrane form by
ic processing (By similarity).

y: Belongs to the tumor necrosis factor family.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL Outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See <http://www.isb-sib.ch/announce/>
ail to license@isb-sib.ch).

1; AAD04375.1; -.

1; C: integral to membrane; ISS.

4; F: CD40 receptor binding; ISS.

10; P: B-cell proliferation; ISS.

4; P: inflammatory response; ISS.

9; P: leukocyte cell adhesion; ISS.

8; P: platelet activation; ISS.

003263; TNF 5.

006052; TNF family.

008983; TNF-like.

003636; TNF_subf.

; TNF; 1.

02; CD40LIGAND.

012; TNF 5; 1.

17; TNF; 1.

251; TNF 1; 1.

049; TNF 2; 1.

unsmembrane; Glycoprotein; Signal-anchor.

1 260

TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

MEMBER 5, MEMBRANE FORM.

12 260

TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

MEMBER 5, SOLUBLE FORM (BY SIMILARITY).

1 22

CYTOPLASMIC (POTENTIAL).

23 46

SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).

47 260

EXTRACELLULAR (POTENTIAL).

111 112

CLEAVAGE (BY SIMILARITY).

177 217

POTENTIAL.

239 239

N-LINKED (GLCNAC...) (POTENTIAL).

50 AA; 28688 MW; 604F69A19E98E70 CRC64;

7.5%; Score 108.5; DB 1; Length 260;

larity 25.5%; Pred.No. 0.16;

Conservative 23; Mismatches 69; Indels 63; Gaps 8;

EEELVAEE--DQDPS-ELNPQTESQDPAPFLNRLVPRRSAPKGRKTRRAIAAH 145

Db 103 EMKKEENIAMQKGDQDPRIAAHVISEASSNPASVL-----RWAPKGYTISNNI

Qy 146 YEYHPRGQDGAQAGVDGTGSGWEEARINSSPLRYNRRIGEFIVTRAGLYLYCC

Db 155 -----ENKQ-----LAVKQGLYYVYAC

Qy 206 EGKAVYLKLDLLVDGVLALRCLEFSAT-----AASSLGFLQLCQVS-----GLI

Db 178 SNRAASQAPP-----VASLCLHSPSGTERVLLRAASSRSGSKPCGQQSILHGGVF

Qy 256 GSSLRIRTLPAWHLKAAPFLTYFGLFQV 283

Db 233 GASVFVNVTDPQSVSHGTGTFTSGLLKL 260

RESULT 5

TNF9 HUMAN

ID TNF9_HUMAN STANDARD; PRT; 254 AA.

AC P41273;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 9 (4-LBB ligand)

DE 1LBB).

GN TNFSF9.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94374434; PubMed=8088337;

RA Alderson M.R., Smith C.A., Tough T.W., Davis-Smith T., Armitage F.

RA Falk B., Roux E., Baker E., Sutherland G.R., Din W.S., Goodwin R.

RT "Molecular and biological characterization of human 4-LBB and its
ligand.";

RL Eur. J. Immunol. 24:2219-2227(1994).

CC -!- FUNCTION: Cytokine that binds to TNFRSF9. Induces the
proliferation of activated peripheral blood T cells. May have
role in activation-induced cell death (AICD). May play a role
in cognate interactions between T cells and B cells/macrophages.

CC -!- SUBUNIT: Homotrimer (Potential).

CC -!- SUBCELLULAR LOCATION: Type II membrane protein.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PLACENTA, LUNG, SKIN
MUSCLE AND KIDNEY.

CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.

CC This SWISS-PROT entry is copyright. It is produced through a col-
laboration between the Swiss Institute of Bioinformatics and the EMBL Out-
station at the European Bioinformatics Institute. There are no restrictions
on its use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for
commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

CC EMBL; U03398; AAA53134.1; -.

CC PIR; I38427; I38427.

CC Genew; HGNC:11939; TNFSF9.

CC MIM; 606182; -.

CC GO; GO:0005102; P: receptor binding; TAS.

CC GO; GO:0006915; P: apoptosis; TAS.

CC GO; GO:0008283; P: cell proliferation; TAS.

CC GO; GO:0007267; P: cell-cell signaling; TAS.

CC GO; GO:0007165; P: signal transduction; TAS.

CC InterPro; IPR006052; TNF family.

CC InterPro; IPR008983; TNF-like.

CC Pfam; PF00229; TNF; 1.

CC SMART; SM00207; TNF; 1.

CC PROSITE; PS00251; TNF 1; 1.

CC PROSITE; PS00049; TNF 2; 1.

KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Polymorphic;
CYTOPLASMIC (POTENTIAL).

DOMAIN 1 28

with the 14-3-3 protein YWAE, MEF2A, MEF2B and MEF2C activity). Interacts with HTATIP and EDNRA. AR LOCATION: Nuclear and cytoplasmic. In the nucleus, it has with distinct subnuclear dot-like structures. Shuttles between the nucleus and the cytoplasm. Treatment with EDN1 results in the nuclear translocation of the protein. The cytoplasm depends on the interaction with the 14-3-3 YWAE and may be due to its phosphorylation. VE PRODUCTS:

alternative splicing; Named isoforms=4;

8WU14-1; Sequence=Displayed;

8WU14-2; Sequence=VSP_007429, VSP_007431;

experimental confirmation available;

8WU14-4; Sequence=VSP_008772;

8WU14-3; Sequence=VSP_007430; experimental confirmation available; the nuclear export sequence mediates the shuttling between the nucleus and the cytoplasm (By similarity).

is phosphorylated by CAK1 (By similarity).

FOUS: Its activity is inhibited by Trichostatin A (TSA), histone deacetylase inhibitor (By similarity).

Y: Belongs to the histone deacetylase family. Subfamily

Ref.1 sequence differs from that shown due to a

t in position 877.

Ref.2 (BAC56929) sequence differs from that shown due to

extension.

OT entry is copyright. It is produced through a collaboration Swiss Institute of Bioinformatics and the EMBL outstation - Bioinformatics Institute. There are no restrictions on its profit institutions as long as its content is in no way this statement is not removed. Usage by and for commercial purposes requires a license agreement (See <http://www.isb-sib.ch/announce/>).

3; AAF63491.1; ALT_FRAME.

2; BAA91474.1; ALT_INIT.

0; BAA91545.1; ALT_INIT.

9; BAB15759.1; ALT_INIT.

1; BAB55363.1; ALT_INIT.

8; BAC56929.1; ALT_SEQ.

7; AAP84704.1; -.

6; -; NOT_ANNOTATED_CDS.

5; CAB55935.1; -.

3; AAH06453.1; ALT_INIT.

5; AAH20505.1; ALT_INIT.

4067; HDAC7A.

7; Cytoplasm; TAS.

8; Histone deacetylase complex; TAS.

4; Cnucleus; TAS.

7; F-histone deacetylase activity; TAS.

6; F-specific transcriptional repressor activity; TAS.

13; F-transcription factor binding; TAS.

13; P-B-cell differentiation; TAS.

13; P-chromatin modification; TAS.

13; P-inflammatory response; TAS.

13; P-negative regulation of myogenesis; TAS.

13; P-neurogenesis; TAS.

14; P-regulation of cell cycle; TAS.

1000286; His_deacetylase.

1; Hist_deacetyl; 1.

270; HDASUPER.

1; nuclear protein; Chromatin regulator;

1; regulation; Repressor; Repeat; Phosphorylation;

splicing.

1 269 TRANSCRIPTION REPRESSION 1 (BY

SIMILARITY).

FT DOMAIN 218 546 TRANSCRIPTION REPRESSION 2 (BY
FT SIMILARITY).
FT DOMAIN 518 865 HISTONE DEACETYLASE.
FT DOMAIN 918 952 NUCLEAR EXPORT (BY SIMILARITY).
FT DOMAIN 1 98 INTERACTION WITH MEF2C (BY SIMILARITY).
FT DOMAIN 49 149 INTERACTION WITH MEF2A (BY SIMILARITY).
FT DOMAIN 877 952 INTERACTION WITH SIN3A (BY SIMILARITY).
FT DOMAIN 197 203 POLY-SER.
FT DOMAIN 368 373 POLY-PRO.
FT ACT_SITE 670 670 BY SIMILARITY.
FT VARSPLIC 1 472 Missing (in isoform 2).
FT VARSPLIC 227 263 /FTID=VSP_007429.
FT VARSPLIC 227 256 /FTID=VSP_008772.
FT VARSPLIC 227 256 Missing (in isoform 4).
FT VARSPLIC 227 256 /FTID=VSP_007430.

Query Match 6.9%; Score 99; DB 1; Length 952;
Best Local Similarity 24.2%; Pred. No. 3.9;
Matches 46; Conservative 21; Mismatches 67; Indels 56; G

QY 11 RRLPLRSLSGRDGGAVR-----QAQPPAPMAAR-----RSQRR
Db 401 RQIPSAEDLETDGGPGQVDDGLEHRELHGQPEARGPAPLQHPQVLLWEQRI
QY 48 -RGEPFGTALLVPLALGLGLALACIGLLLVVSLGRSLASQAEPQAEELVAEDQI
Db 461 PRGSTGDTVLLPLAQGHRLPLS-----RAQSSPAAPASLSAPEASQARVLSSET
QY 107 NPQT-----EESQDP-----APFLNLRVRRSPAPKGRK
Db 516 LPFTTGLIYDSVLMKHQCSGDNRSRHPHAGRIQSIWSRLQERGLRSQCECLGRKF
QY 140 RAIAAHVEVH 149
Db 576 ELQSVHSEH 585

RESULT 7

CG22 ANTMA STANDARD; PRT; 441 AA.
AC P34801;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G2/mitotic-specific cyclin 2.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; aster
OC Lamiales; Lamiales; Antirrhinaceae; Antirrhineae; Antirrhinum.
OX NCBI_TaxID=4151;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94148008; PubMed=8313906;
RA Robert P.R.; Coen E.S.; Murphy G.J.P.; Doonan J.H.;
RT "Patterns of cell division revealed by transcriptional regulator
genes during the cell cycle in plants.";
RL EMBO J. 13:616-624(1994).
CC -!- FUNCTION: Essential for the control of the cell cycle at the
CC (mitosis) transition. G2/M cyclins accumulate steadily during
CC and are abruptly destroyed at mitosis.
CC -!- SUBUNIT: Interacts with the CDC2 and CDK2 protein kinases to
CC a serine/threonine kinase holoenzyme complex. The cyclin sub
CC imparts substrate specificity to the complex.
CC -!- DEVELOPMENTAL STAGE: Accumulates steadily during G2 and is
CC abruptly destroyed at mitosis.
CC -!- SIMILARITY: Belongs to the cyclin family. Cyclin A8 subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a col
CC between the Swiss Institute of Bioinformatics and the EMBL out
CC the European Bioinformatics Institute. There are no restriction
CC by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for (

res a license agreement (See <http://www.isb-sib.ch/announce/> to license@isb-sib.ch).

CAAS3729.1; -

41710.

1VIN.

06670; Cyclin.

04367; Cyclin Cterm.

06671; Cyclin_N.

cyclin; 1.

cyclin; 1.

; CYCLIN; 2.

92; CYCLIN; 1.

ycle; Cell division; Mitosis.

AA; 49205 MW, E6E4C037C98880A7 CRC64;

6.7%; Score 97; DB 1; Length 441;

urity 24.0%; Pred. No. 2.3;

nservative 44; Mismatches 104; Indels 74; Gaps 15;

QAPPMAARRSQR-----RGRGPGTALVPLALGLALACLGLL 73

XS-----MAVEKKNRAGDIGNVTVRGEGKALPQVSRPIIRGP-----CAQLI 69

-----LAVSLGRASLS-----AQEPAQEELVAEEDQDPSELNPQTEESQ 114

AAENNNKSLAVNAGADGALPIKEAVARVPQKTKVSKQEIIIEISPDTEKKX 129

LNRLVRRRS-----APGRKTRARRAIAHVEVHPRPGQ-----DGAAGVGDGV 165

LEKEITGSKLKKKAPTTLTSTLTARSKAASV-VRTKPKQEIQVIDAADVNDLAV 186

-----ARINSSPLRY-----NRQIGFIVTRAGLYLYLCVHFDP-----EGKAVYL 212

DMYFKYKAENDSRPHDYMDSQPEINEKK-----RAILLIDLWLQVHYKFAELSPETLYL 244

ADGVILALRC-----LEEFATAASSLGPQLRLCQVSGLLALRPGS 257

VDRYLASKTSRRELQLLGMSSMLTASXYEEIWAPEVNDLVCIISDGS 295

STANDARD; PRT; 201 AA.

rel. 40, Created)

rel. 40, Last sequence update)

rel. 42, Last annotation update)

pha precursor (LT-alpha) (TNF-beta) (Tumor necrosis superfamily member 1).

OR TNFB.

lii (Tamar wallaby).

azoa; Chordata; Craniata; Vertebrata; Euteleostomi; theria; Diprotodontia; Macropodidae; Macropus.

5;

N.A.

48; PubMed=10826697;

Deane E.M.;

of lymphotoxin alpha (LT-alpha) from a marsupial, lii,"

9-403(2000).

Cytokine that in its homotrimeric form binds to TNFR1, TNFRSF1B/TNFR and TNFRSF14/HVEM. In its meric form with LTB binds to TNFRSF3/LTBR. Lymphotoxin is y lymphocytes and cytotoxic for a wide range of tumor itro and in vivo.

homotrimer, and heterotrimer of either two LTB and one its or (less prevalent) two LTA and one LTB subunits (By /).

AR LOCATION: Secreted (homotrimer) and membrane i (heterotrimers) (By similarity).

This SWISS-PROT entry is copyright. It is produced through a colla between the Swiss Institute of Bioinformatics and the EMBL out the European Bioinformatics Institute. There are no restrictions use by non-profit institutions as long as its content is in CC modified and this statement is not removed. Usage by and for CC entities requires a license agreement (See <http://www.isb-sib.ch/> or send an email to license@isb-sib.ch).

EMBL; AF119336; RAD41773.1; -

HSSP; P01374; LTR.

InterPro; IPR006053; TNF abc.

InterPro; IPR006052; TNF family.

InterPro; IPR008983; TNF like.

InterPro; IPR003636; TNF_subf.

Pfam; PF00229; TNF; 1.

PRINTS; P01234; TNCRSISFCT.

ProDom; PD002012; TNF_subf; 1.

SMART; SM0207; TNF; 1.

PROSITE; PS00251; TNF 1; 1.

PROSITE; PS0049; TNF 2; 1.

Cytokine; Glycoprotein; Signal.

FT SIGNAL 1 27 BY SIMILARITY.

FT CHAIN 28 201 LYMPHOTOXIN-ALPHA.

FT CARBOHYD 93 93 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 201 AA; 21536 MW; 8C4C371CB5091627 CRC64;

Query Match 6.6%; Score 95.5; DB 1; Length 201;

Best Local Similarity 23.2%; Pred. No. 1.2;

Matches 44; Conservative 27; Mismatches 84; Indels 35; G

QY 107 NPQTESQDPAPFLNRLVRRSAPKGRKTRARRAI--AAHYEVHPRPGDGAQAG

DB 30 NPDNHHSSSPAP-----PQTAQHLSQKSLKEITLKPAHL-----VGDPSVQDS.

QY 165 VSGWEARINSSP-LRYNRQI--GEFIVTRAGLYLYLCVHFDEKA-----V

DB 76 --W---RANTDHAFLRHGFSLSNNLSLVPTSGLYEVSQVWFSGASCSEITPTLL

QY 215 DLLVDG---VLALRCLEEFSAATAASSLGPQLRLCQVSGLLALRPGSSLIRTLPMW

DB 130 EVLLFSSKYQHVPLLSAOKSVCSGTQGPWMSVYQCAVFLITQGRSLTYTIDGVS

QY 272 APFLTYFGLF 281

DB 190 SPSSVFFGAF 199

RESULT 9

TNF6_MACMU

ID TNF6_MACMU STANDARD; PRT; 280 AA.

AC Q9MYL6; Q9BDM5;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen li (CD95L protein).

DE TNFSF6 OR FASL OR CD95L.

GN Macaca mulatta (Rhesus macaque),

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), an

OS Macaca nemestrina (Pig-tailed macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopitheciinae; Macaca.

OC NCBI_TaxID=9544, 9541, 9545;

OX [1]

RN SEQUENCE FROM N.A.

RP SPECIES=Mulatta; TISSUE=Lymphocytes;

RX MEDLINE=21383618; PubMed=11491535;

RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P., Weiss W.R., Ansari A.A.;

RA "Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules.";

RT Immunogenetics 53:315-328(2001).

RL

79 GLCLLVFFWVVALVGLG-LGMFQLFHLQKEL-----AELRESTSQKHTAA
 122 KLVPRRSAPKGRKTRARRAIAAHYEHPRPGDQAQGVDTGSGWEAA-RINSS
 129 QIGHP--SPPEKKEQRK-VAHLTKGNSRSMPL-:::-----WEDTYGIVLL
 181 YNRQIGFEIVTRAGLYLYICQVHFDEGKA-----VYLKLD-----LLVDGV
 175 YKK--GSLVINETGLYFVSKVYF-RGQCTNLPLSHKYMRNSKYPODLVYMGSG
 226 CLEEFSAATASSIGPOLRQCQVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYFGFL
 232 CTTGQMAHSSYLGAENVLTSADHLY-----VNVSELSLVNFEESQ--TFFGLY

RESULT 10
 RHO MICLU
 ID _RHO MICLU STANDARD; PRT; 690 AA.
 AC P52154;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcription termination factor rho.
 DE RHO.
 OS Micrococcus luteus (Micrococcus lysodeikticus).
 OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococccineae; Micrococcaceae; Micrococcus.
 ON NCBI_TaxID=1270;
 RN [1]_SEQUENCE FROM N.A., AND SEQUENCE OF 1-5 AND 289-297.
 RP STRAIN=EM;
 RC MEDLINE=96132802; PubMed=8557681;
 RX Nowatzke W.L., Richardson J.P.;
 RA "Characterization of an unusual Rho factor from the high G + C gr:
 RT Positive bacterium Micrococcus luteus.";
 RT J. Biol. Chem. 271:742-747(1996).
 RN [2]
 RP SEQUENCE OF 205-690 FROM N.A.
 RC STRAIN=EM;
 RX MEDLINE=94327472; PubMed=8051015;
 RA Opperman T., Richardson J.P.;
 RT "Phylogenetic analysis of sequences from diverse bacteria with
 RT homology to the Escherichia coli rho gene.";
 RT J. Bacteriol. 176:5033-5043(1994).
 RN [3]
 RP REVISION TO 500.
 RC STRAIN=EM;
 RX Nowatzke W.L.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM
 CC THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF I
 CC RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM A
 CC DNA TEMPLATE. RNA-DEPENDENT NTPASE WHICH UTILIZES ALL FOUR
 CC RIBONUCLEOSIDE TRIPHOSPHATES AS WELL AS DATP AS SUBSTRATES, I
 CC HAS A SIGNIFICANT LOWER ACTIVITY WITH CTP.
 CC -!- SUBUNIT: Homohexamer (By similarity).
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a col
 CC between the Swiss Institute of Bioinformatics and the EMBL out
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is not
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch>).
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L27277; AAE18671.1; -.
 CC HSSP; P03002; 1A63.
 CC InterPro; IPR003593; AAA_ATPase.
 CC InterPro; IPR000194; ATPase_a/bcentre.
 CC InterPro; IPR002059; Cold_shock.
 CC InterPro; IPR008994; Nucleic_acid_OB.
 CC InterPro; IPR004665; Term_rho.

ATP-synt_ab; 1.
AAA; 1.
CSP; 1.
00767; rho; 1.
ermination; Helicase; ATP-binding; RNA-binding.

0
0 35 RNA-BINDING (RN2) (BY SIMILARITY).
3 321 RNA-BINDING (RN1) (BY SIMILARITY).
1 448 ATP (POTENTIAL).
1 291 G -> P (IN REF. 2; AA SEQUENCE).
AA; 75030 MW; F7704C75EE1B8998 CRC64;
6.5%; Score 93.5; DB 1; Length 690;
rity 25.9%; Pred.No.7.3; Indels 41; Gaps 7;
nservative 19; Mismatches 63; Indels 41; Gaps 7;

ET SARLPLPRSLGRDGGAVRQAPPAARRSRRRGPGTALLVPLAL 61
DAERAQAAPAETAEPAAASSDDAAP-AAERPARRSRRADTS--APAAA 111
LA LGLLLAVSLGSRASLSAQEAELVAEDDPSELNPQTESODPAPFLN 121
QA-----EAREAQEAPRE---FASDQRSGSEARDEGED----- 150
RRSAFKGRKTRARRAIAAHYEVHPPGQGAQAGVDGVTVSG 167
QSE-----RRSRGRR-----RAGDDDAQQGQDRSDG 179

STANDARD; PRT; 310 AA.

el. 34, Created)
el. 34, Last sequence update)
el. 42, Last annotation update)
rotein RVC497/MT0517/MB0508.
17 OR MTCY20G9.23 OR MB0508.
tuberculosis, and
bovis.
nobacteria; Actinobacteridae; Actinomycetales;
eae; Mycobacteriaceae; Mycobacterium.
3, 1765;

N.A.
rculosis; STRAIN=H37Rv;
87; PubMed=9634230;
sch R., Parkhill J., Garnier T., Churche C., Harris D.,
iglmeier K., Gas S., Barry C.E. III, Tekala F.,
sham D., Brown D., Chillingworth T., Connor R.,
llin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
gels K., Krogh A., McLean J.J., Moule S., Murphy L.,
orne J., Quail M.A., Rajandream M.A., Rogers J.,
ger K., Skelton S., Squares S., Squares R.,
Taylor K., Whitehead S., Barrrell B.G.;
he biology of Mycobacterium tuberculosis from the
e sequence.";
-544(1998).

N.A.
rculosis; STRAIN=CDC 1551 / Oshkosh;
94; PubMed=12218036;
D., Alland D., Eisen J.A., Carpenter L., White O.,
eBoy R., Dodson R., Winn M., Haft D., Hickey E.,
Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
terback T., Weidman J., Khouri H., Gill J., Mikula A.,
obs W.R. Jr., Venter J.C., Fraser C.M.;
comparison of Mycobacterium tuberculosis clinical and
ains";
184:5479-5490(2002).

N.A.
N.A.; STRAIN=AF2122/97;

RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoisier S., Grondin S., Iacox C., Monsemp C., Simon S.
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis."
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC CC -!- SIMILARITY: TO M.LEPRAE ML2433.
CC This SWISS-PROT entry is copyright. It is produced through a colla
CC between the Swiss Institute of Bioinformatics and the EMBL out
CC the European Bioinformatics Institute. There are no restrictions
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for co
CC entities requires a license agreement (See http://www.isb-sib.ch/a
CC or send an email to license@isb-sib.ch).
CC EMBL; Z77162; CAB00923.1; --
DR EMBL; AE006952; AAK44740.1; --
DR EMBL; BX248335; CAD93371.1; --
DR PIR; D70745; D70745.
DR TIGR; MT0517;
DR Tuberculis; Rv0497; --
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 231 251 POTENTIAL.
FT TRANSMEM 257 277 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT DOMAIN 33 39 POLY-ARG.
FT DOMAIN 197 202 POLY-ALA.
SQ SEQUENCE 310 AA; 33092 MW; 4954027F694DF5C2 CRC64;
Query Match 6.4%; Score 93; DB 1; Length 310;
Best Local Similarity 24.4%; Pred. No. 3.2;
Matches 80; Conservative 25; Mismatches 103; Indels 120; Ga
QY 14 PLPSRLGSRDGA----VRQAPPAPMAARSQRERG-----
DB 6 PETESSGNRQISVAELLARQGVGTGAP--ARRRRRRGSDAITVAELTGTEPIIRDI
QY 48 RCEPTGALLVPALGLGLALACLGILLAVSLGSRASLSAQEAELVAE-----
DB 64 AGPDHAHSQS PANGR-----VQGEAAPQSPAEFVAEQ-VASEPTFT
QY 100 DDPSLNPQTESODPAPFLNLVPRR-----SAPKRKTFRARRI--AAHY--
DB 110 QPEWRPKSPQDRRESGPSELSEYPRLRHTTHSDRAPAGPPSGAHEMGDPVVEHYPI
QY 147 -----EVHPRQDG-----AQAGVDGVTVSGWEHARINSSS-----PI
DB 170 DVLDTEVGEEATEVREAQPGRGERHAAAAAGTDVEGDGAARVARRALDVVVPI
QY 183 ROIGEFIVTR-----AGLYLYCQVFIDE-----GKAVYLKLIDL-----LVDCGV
DB 230 ---GALVLVQLSILAFAFGNGLF-----TAFDLWFWNSIVALVLSVMILGLVSVS
QY 226 CLEFSATS-----AASSLGPQLRLCQ 246
DB 282 KTEDIASTLIJAVAVGALITLGP-LALLQ 308
RESULT 12
TN14 HUMAN STANDARD; PRT; 240 AA.
ID TN14 HUMAN STANDARD; PRT; 240 AA.
AC Q4357; O75476; OSWF8; Q96LD2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 14 (Herpesvirus ei
DE mediator-ligand) (HVEM-L).
GN TNFSF14 OR LIGHT OR HVEM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

heria; Primates; Catarrhini; Hominidae; Homo.
06;
(N.A. (ISOFORM 1)).
340; PubMed=9462508;
bner R., Montgomery R.I., Koche K.D., Cheung T.C.,
en S., Murphy M., Eisenberg R.J., Cohen G.H., Spear P.G.,
' member of the TNF superfamily, and lymphotoxin alpha are
erpesvirus entry mediator.";
-30(1998).
(N.A. (ISOFORM 1)), AND CHARACTERIZATION.
532; PubMed=9765287;
McDonnell P.C., Brigham-Burke M., Lyn S.D., Minton J.,
le K., Spanpanato J., Silverman C., Hensley P.,
Emery J.G., Deen K., Eichman C., Chabot-Fletcher M.,
ung P.R.;
entry mediator ligand (HVEM-L), a novel ligand for
mulates proliferation of T cells and inhibits HT29 cell
l. 273:27548-27556(1998).
(N.A. (ISOFORM 2)), AND PROCESSING.
948; PubMed=11673523;
Butrovich K.D., Houshmand P., Edwards W.R., Ware C.F.;
acterization of LIGHT reveals linkage to an immune
is on chromosome 19p13.3 and distinct isoforms generated
splicing or proteolysis.";
.67:5122-5128(2001).
(N.A.
1257; PubMed=12477932;
L., Feingold E.A., Grouse L.H., Derge J.G.,
Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
Zeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Marusina K., Farmer A., Rubin G.M., Hong L.,
Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Quellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Mazny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
ton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
ladan A., Young A.C., Shevchenko Y., Bouffard G.G.,
l., Touchman J.W., Green E.D., Dickson M.C.,
S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
Schein J.E., Jones S.J.M., Marra M.A.;
and initial analysis of more than 15,000 full-length
cDNA sequences.";
acad. Sci. U.S.A. 99:16899-16903(2002).
Cytokine that binds to TNFRSF3/LTR. Binding to the
ceptor TNFRSF6B modulates its effects. Activates NFkB,
as the proliferation of T cells, and inhibits growth of
xarcinoma HT-29. Acts as a receptor for Herpes simplex
Homotrimer.
AR LOCATION: Type II membrane protein and secreted
1); Cytoplasmic (isoform 2).
[VE PRODUCTS:
ernative splicing; Named isoforms=2;
X3557-1; Sequence-Displayed;
ynonyms=LIGHT delta-TW;
X3557-2; Sequence=VSP_006452;
PECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT ALSO
THE BRAIN. WEAKLY EXPRESSED IN PERIPHERAL LYMPHOID
AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KIDNEY,
(PRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OR

CC NONHEMATOPOIETIC TUMOR LINES.
CC -!- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.
CC -!- PTM: N-glycosylated.
CC -!- PTM: The soluble form of isoform 1 derives from the membrane
CC by proteolytic processing.
CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC -!- CAUTION: Ref.4 sequence differs from that shown due to a
CC frameshift in position 178.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a coll
CC between the Swiss Institute of Bioinformatics and the EMBL out
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is ir
CC modified and this statement is not removed. Usage by and for
CC entities requires a license agreement (See <http://www.isb-sib.ch/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF036581; AAC39563.1; -.
CC EMBL; AF064090; AAC25169.1; -.
CC EMBL; AY028261; AAK26160.1; -.
CC EMBL; BC018058; AAH18058.1; ALT_FRAME.
CC HSSP; P01375; 4TSV.
CC Genew; HGNC:11930; TNFSF14.
CC MIM; 604520; -.
CC GO; GO:0005102; P:receptor binding; TAS.
CC GO; GO:0006917; P:induction of apoptosis; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR006053; TNF_abc.
CC InterPro; IPR006052; TNF_family.
CC InterPro; IPR008983; TNF_like.
CC Pfam; PF00229; TNF; 1.
CC PRINTS; PR01234; TNFCROSISFCT.
CC ProDom; PD02012; TNF_subf; 1.
CC SMART; SM00207; TNF; 1.
CC PROSITE; PS00251; TNF_1; FALSE_NEG.
CC PROSITE; PS00049; TNF_2; 1.
CC Cytokine; Transmembrane; Glycoprotein; Signal-anchor;
KW Alternative splicing.
FT CHAIN 1 240 TUMOR NECROSIS FACTOR LIGAND SUPERF
FT CHAIN 793 240 MEMBER 14, MEMBRANE FORM.
FT CHAIN TUMOR NECROSIS FACTOR LIGAND SUPERF
FT DOMAIN 1 37 MEMBER 14, SOLUBLE FORM.
FT TRANSMEM 38 58 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 59 240 SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO
FT SITE 82 83 (POTENTIAL).
FT SITE 82 83 EXTRACELLULAR (POTENTIAL).
FT DISULFID 154 187 CLEAVAGE (POTENTIAL).
FT CARBOHYD 102 102 POTENTIAL.
FT VARSPPLIC 38 73 N-LINKED (GLCNAC. . .).
FT CONFLICT 120 120 Missing (in isoform 2).
FT CONFLICT 214 214 /FTid=VSP_006452.
FT CONFLICT 214 214 L -> V (IN REF. 4).
FT CONFLICT 214 214 E -> K (IN REF. 2).
SQ SEQUENCE 240 AA; 26351 MW; 49D0BF67E1390B39 CRC64;
Query Match 6.4%; Score 92; DB 1; Length 240;
Best Local Similarity 23.7%; Pred. No. 2.8;
Matches 44; Conservative 19; Mismatches 57; Indels 66; C
QY 29 QAQPAPMAARSQRRCRGPCTALLVPLALGLGLALACGLLLAVVLSGRK
DB 16 QTDPTFTLGRSHRRSCSVARVGLGLLL-LIMGAGLAVQGVFLQLHWRLG-
QY 89 EPAQELVAEEDQDPSELNPOTESQDPAPFLNRLVRRPRRSPKGRKTRARRAI
DB 67 -----EMV-----TRLPDGPAGSWEQLIQERS-----
QY 149 HPRPQDGAQGVDTGTVSGWEAKINSSPLRYNRQI-----GEFIVT
DB 93 NPAHLTGANSLSLTG-----SCGPLLWETQLGLAFLRGLSYHDGALVVT
QY 197 YLYQCV 202

1 V 147
FT DOMAIN 4 69 PRO-RICH.
FT SITE 128 64 POLY-PRO.
FT DISULFID 201 232 CLEAVAGE (BY SIMILARITY).
FT CARBOHYD 183 183 POTENTIAL.
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 280 AA; 31407 MW; 729EA60067B7D398 CRC64;
Query Match 6.4%; Score 92; DB 1; Length 280;
Best Local Similarity 20.8%; Pred. No. 3.4; Indels 74; Ga
Matches 62; Conservative 44; Mismatches 118;
QY 13 LPPLPSLSRGGAVRQAPAPMAARRSQRRRGRGPGTALLVPLAL-----
Db 30 LPCPTSVRRPQORPPPPPPPLP-----PPPPPLPLPLPKKSGN
QY 62 GLGLALACILGLLVAVLSIGSRASLSAQPAQELVAEEDQDPSSELNPQTESQDPAE
Db 79 GLCLLVMPFMVLVALVGLG--LGMFLQLFLQKEL-----AELRESTSQKHTASE
QY 122 RLVRPRRSPKGRKTRARRATAAHYEVHPRQDGAQAGVDGTVSGWEEA-RINSSE
Db 129 QIGHP---SPPEKKEQRK--VAHUTGPNRSKSMPL-----WEDTYGIVLLK
QY 181 YNRQIGEFIVTRAGLYLYCYQVHFDEGKA-----VYLKLD-----LLVDGVI
Db 175 YKK--GGLVINETGLYFYVSKYVF-RGQSCNPLPLSHKVMYMSKYPQDLVMEGKO
QY 226 CLUEPSATAASLSGLQLRLCQVSGLLAURPGSLRIRTLFWAHLKAAPELTYFGLF
Db 232 CTTGQWHAHSSYLGAVFNLSTDLHY-----VNVSELSLVNFESQ--TFGLYI
RESULT 14
NFC4 HUMAN
ID NFC4 HUMAN STANDARD; PRT; 902 AA.
AC Q14934;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Nuclear factor of activated T-cells, cytoplasmic 4 (T cell
transcription factor NFAT3) (NF-ATc4) (NF-AT3).
GN NFATC4 OR NFAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=95269130; PubMed=7749981;
RA Hoey T., Sun Y.-L., Williamson K., Xu X.;
RT "Isolation of two new members of the NF-AT gene family and function
characterization of the NF-AT proteins.";
RL Immunity 2:461-472(1995).
RN [2]
RP REVIEW.
RX MEDLINE=99189746; PubMed=10089876;
RA Crabtree G.R.;
RT "Generic signals and specific outcomes: signaling through Ca2+,
calcineurin, and NF-AT.";
RL Cell 96:611-614(1999).
CC -!- FUNCTION: Plays a role in the inducible expression of cytokine
genes in T cells, especially in the induction of the IL-2 and
IFN-gamma.
CC -!- SUBUNIT: Member of the multicomponent NFATC transcription com
plex that consists of at least two components, a pre-existing
cytoplasmic component NFATC2 and an inducible nuclear component
NFATC1. Other members such as NFATC4, NFATC3 or members of the
activating protein-1 family, MAP, GATA4 and Cbp/p300 can also
form the complex. NFATC proteins bind to DNA as monomers.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic for the phosphorylated form

ter activation that is controlled by calcineurin-phosphorylation. Rapid nuclear exit of NFATC is thought mechanism by which cells distinguish between sustained and transient calcium signals. The subcellular localization of NFATC is a key role in the gene transcription. NFATC1 is highly expressed in placenta, lung, kidney, and ovary. Weakly expressed in spleen and thymus. Not in peripheral blood lymphocytes. NFATC1 allows DNA-binding and NFATC1 domain (NSD) allows DNA-binding and NFATC1 interacts with API factors (By similarity). NFATC1 is phosphorylated by NFATC1-kinase; dephosphorylated by NFATC1 phosphatase (By similarity). NFATC1 belongs to the Rel/Dorsal family.

OT entry is copyright. It is produced through a collaboration Swiss Institute of Bioinformatics and the EMBL outstation - Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

AA79175.1; -
462; -
778; NFATC4.
-
3; P:transcription co-activator activity; TAS.
4; P:inflammatory response; TAS.
6; P:transcription from Pol II promoter; TAS.
007110; Ig-like.
002909; IPT TIG.
00451; NF Rel dor.
008366; NFAT.
008967; P53-like.
; RHD; 1.
; TIG; 1.
89; NUCFACTORATC.
9; IPT; 1.
204; REL_1; FALSE NEG.
254; REL_2; 1.
regulation; Activator; Nuclear protein; DNA-binding; phosphorylation.
62 69
POLY-PRO.
14 119 CALCINEURIN-BINDING.
13 293 2 APPROXIMATE SP REPEATS.
13 229 SP 1.
77 293 SP 2 (APPROXIMATE).
97 304 POLY-PRO.
68 270 NUCLEAR LOCALIZATION SIGNAL.
30 437 NUCLEAR BINDING.
72 674 NUCLEAR LOCALIZATION SIGNAL.
2 AA; 95472 MW; 559F15F7647A47C6 CRC64;
6.4%; Score 92; DB 1; Length 902;
arity 25.0%; Pred. No. 13;
conservative 15; Mismatches 73; Indels 56; Gaps 8;
SLGSDGAVTQAQPPAPMAARRS-----QRRGRGRGPGTALLVPLGLALAC 69
LQGGEDSLLLSAGPTPASPAPSPGCKRYSSSGTPSSA-----SPALSR 286
LAVSLGSRASLSAQEPALVAEDDQPSLNPQTEESQDPAPFLNRLVPRRS 129
-----SLGEGS-----EPPPPPL-PLARDPGSPGFEDVVGAPPAES 325
GRKTRARRAIA-----AHVEVHPQDGAQAGVD-----GTVSG 167
TERTTSSEQAVLPRSEEPASCNGLPLGAEESVAPPGGSRKEVAGMDYLAVPSPLA 385
LRNSSPL 179
RIGHSPI 397

RESULT 15

MLH1_MOUSE STANDARD; PRT; 760 AA.
ID MLH1_MOUSE
AC Q9UK91; Q62454;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA mismatch repair protein Mlh1 (MutL protein homolog 1).
GN MLH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Kumaran M., Rao M.R.S.;
RT "Cloning of the cDNA of the MutL homolog, MLH1 from mouse testis.
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-151 FROM N.A.
RX MEDLINE=96270514; PubMed=8674118;
RA Edelmann W., Cohen P.E., Kane M., Lau K., Morrow B., Bennett S.,
RA Udelmann W., Kunkel T., Cattoretti G., Chaganti R., Pollard J.W.,
RA Kolodner R.D., Kucherlapati R.;
RT "Meiotic pachytene arrest in MLH1-deficient mice.";
RL Cell 85:1125-1134(1996).
CC -!- FUNCTION: Probably involved in the repair of mismatches in DN
CC -!- SUBUNIT: HETERODIMER OF MLH1 AND PMS2 OR MLH1 AND MLH3 (BY
CC SIMILARITY).
CC -!- SIMILARITY: Belongs to the DNA mismatch repair mutL/hexB fami
CC This SWISS-PROT entry is copyright. It is produced through a coll
CC between the Swiss Institute of Bioinformatics and the EMBL out
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for c
CC entities requires a license agreement (See <http://www.isb-sib.ch/>
CC or send an email to license@isb-sib.ch).

EMBL; AF250844; AAC52672.1; -
EMBL; U60872; AAC52672.1; -
EMBL; U59881; AAC52672.1; JOINED.
EMBL; U59882; AAC52672.1; JOINED.
EMBL; U59883; AAC52672.1; JOINED.
EMBL; U59884; AAC52672.1; JOINED.
HSP; P23367; 1BKN.
MGD; MGI:101938; Mhl1.
GO; GO:0000793; C:condensed chromosome; IDA.
GO; GO:0007060; P:meiotic chromosome segregation; IMP.
GO; GO:0007126; P:meiosis; IDA.
GO; GO:0007131; P:meiotic recombination; IMP.
GO; GO:0006298; P:mismatch repair; IDA.
InterPro; IPR002099; ATPbind ATPase.
Pfam; PF01119; DNA_mis_repair; 1.
Pfam; PF02518; HATPase_C; 1.
SMART; SM00387; HATPase_C; 1.
TIGRFAMs; TIGR00585; mutL; 1.
PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
KW DNA repair.
SQ SEQUENCE 760 AA; 84679 MW; 173C809372A29186 CRC64;
Query Match 6.3%; Score 90.5; DB 1; Length 760;
Best Local Similarity 22.7%; Pred. No. 14;
Matches 70; Conservative 31; Mismatches 126; Indels 81; G

QY 13 LPLPSRLSGRGGAVRO-----AQPAPMAARRSQRGRGPGTALLVPLA
DB 436 LPAPAAAESENLERESLMETSDAAKQAFTSPGSRKRRH--EDSDVEMVNF
QY 65 LALAC-----LGLLAVVSLGSRASLSAQEPALVAEDDQPSLNPQ--TEES

06:25:24 2004

us-09-245-198a-4.rsp

1

YPRRIINLTSVLQEEISERCHETLRB--ILRNHSFVGCVPQWALAQHOTKL 551
ILVPRRSAPKGRKTRARRAIAHYEV-----HPRPCDGAQAGVDG 163
-----TTKJSEELFYQILIYDFANFGVLRRLSEPAFLDLAMLALDS 596
NEEARINSSPLRYNQIGEFIVTRAGLYLYGCVHFDEGKAVYLKLDLIVDG-- 220
WEDDGPKEGLABY---IVEFLAKKAEMLADYESVEIDE--EGNLIIGLPLIDSYP 652
--VIALR-----CLEEFSATAAS--SLGPQ--LRLCQVSGLLALRPSS 258
JPIFILRLATEVNWDEEKECFESLSKECAMFYSIKQYILEESTLSGQQSDMEGST 712
CLEW 266
--FW 716

April 7, 2004, 17:45:19
secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

search, using sw model

il 7, 2004, 17:37:32 ; Search time 56 9116 Seconds
(without alignments)
1409.967 Million cell updates/sec

09-245-198A-4
SLLDFFISARRLPLPRSLG.....PWAHLKAAPLTIFGLFQVH 284

SUM62

op 10.0 , Gapext 0.5

6107 seqs, 282547505 residues

s satisfying chosen parameters: 1586107

th: 0

th: 2000000000

nimum Match 0%

ximum Match 100%

sting first 45 summaries

Geneseq 29Jan04: *

geneseq1980a: *

geneseq1990a: *

geneseq2000a: *

geneseq2001a: *

geneseq2002a: *

geneseq2003a: *

geneseq2003bs: *

geneseq2004a: *

the number of results predicted by chance to have a
than or equal to the score of the result being printed,
d by analysis of the total score distribution.

SUMMARIES

ry	ch	Length	DB	ID	Description
.0	284	2	AaW47525	AaW47525 Homo sapi	
.8	249	2	AaY09369	AaY09369 Human tum	
.8	249	3	AaY95338	AaY95338 Human PRO	
.8	249	3	AaB07526	AaB07526 Amino aci	
.8	249	5	AaU86129	AaU86129 Human PRO	
.8	249	6	AaB42315	AaB42315 Human TWB	
.8	249	7	AdC35206	AdC35206 Human TNF	
.6	249	2	AaW29745	AaW29745 TNF relat	
.6	249	4	AaE00891	AaE00891 Human TRE	
.6	249	7	AdC97712	AdC97712 Murine FL	
.8	273	4	AaU03499	AaU03499 TWEAK ext	
.5	208	2	AaW93590	AaW93590 Human TNR	
.6	225	2	AaW47524	AaW47524 Mus muscu	
.6	225	3	AaB07527	AaB07527 Amino aci	
.0	211	2	AaW93591	AaW93591 Mouse TNR	
.8	189	2	AaW29746	AaW29746 TNF relat	
.8	189	4	AaE00892	AaE00892 Human U14	
.7	146	4	AaE00895	AaE00895 Human TRE	
.0	325	4	AaB67553	AaB67553 Drosophil	
.0	409	5	AaU77718	AaU77718 Drosophil	
.5	211	3	AaY58216	AaY58216 Canine na	
.5	260	3	AaY58215	AaY58215 Canine CD	
.4	254	4	AaR64190	AaR64190 Human 4-1	
.4	254	2	AaW26657	AaW26657 Human 4-1	
.4	254	5	Abb75953	Abb75953 Human cyt	

26	106.5	7.4	254	6	ABR42312	Hum
27	106.5	7.4	254	7	ADC35200	Hum
28	106.5	7.4	254	7	ADD18780	Hum
29	106	7.3	1428	3	AA97033	Cas
30	105.5	7.3	406	5	AAU77717	Dro
31	104.5	7.2	779	5	ABB07845	Hum
32	104	7.2	409	5	AAU77716	Dro
33	104	7.2	1323	2	AAR55248	N-m
34	102	7.1	256	4	AAM25657	Hum
35	100.5	7.0	647	2	AAW04327	Rat
36	100	6.9	220	4	AA62340	Gp1
37	99	6.9	574	3	AA97032	Cas
38	99	6.9	855	7	AD96563	Hum
39	99	6.9	915	6	ABP56824	Hum
40	99	6.9	1008	4	AAW78891	Hum
41	99	6.9	1020	4	AAW79875	Hum
42	97.5	6.8	1560	7	ADE71264	Nov
43	96.5	6.7	1097	4	ABG25655	Nov
44	96.5	6.7	1631	4	ABG22481	Nov
45	96.5	6.7	19938	6	ABP76681	Str

ALIGNMENTS

RESULT 1
AAW47525
ID AAW47525 standard; protein; 284 AA.
XX
AC AAW47525;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens tumour necrosis factor related ligand (TRELL).
XX
KW TRELL; tumour necrosis factor related ligand; tnfr; treatment; can
autoimmune disease; immune system; stimulation; suppression;
graft rejection.
XX
OS Homo sapiens.
XX
PN WO9805783-Al.
XX
PD 12-FEB-1998.
XX
PF 07-AUG-1997; 97WO-US013945.
XX
PR 07-AUG-1996; 96US-0023541P.
18-OCT-1996; 96US-0028515P.
18-MAR-1997; 97US-0040820P.
XX
PA (BIOJ) BIOGEN INC.
PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
PI Chicheportiche Y, Browning JL;
DR WPI: 1998-145619/13.
DR N-PSDB; AAV18600.
XX
PT Tumour necrosis factor related ligand - useful for, e.g. treating
auto-immune disease and immune responses to tissue grafts.
XX
PS Claim 12; Page 50-51; 69pp; English.
XX
CC The sequence is that of human tumour necrosis factor related ligand.
(TRELL). TRELL or active fragments can be included with a carrier
pharmaceutical compositions to treat cancer, autoimmune diseases c
immune responses to tissue grafts, or to stimulate or suppress the
system. It is useful to screen for TRELL receptors, by labelling a
detectable label and screening compositions for binding. Agents
interfering with TRELL-receptor binding can also be screened for,
then be administered, optionally with interferon- gamma, to induce
death or treat, suppress or alter immune responses (especially inv

ccinoma cells) involving a signal pathway between TRELL and its coding sequence can be used in gene therapy for TRELL-
 iters in mammals (especially humans), e.g. tumours,
 i inflammatory diseases or inherited genetic disorders, by
 to cells, and expressing, therapeutically effective amounts
 s.g. a virus comprising a gene encoding TRELL. It may also
 the preparation of prepare probes for screening
 etic DNAs for TRELL-encoding sequences and for antisense

AA;
 100.0%; Score 1444; DB 2; Length 284;
 arity 100.0%; Pred. No. 2.4e-129;
 onservative 0; Mismatches 0; Indels 0; Gaps 0;
 DFEISARRLPLPSLGRSDGAVTQAPPPAPMAARRSQRGRGEGTALLVPLA 60
 DFEISARRLPLPSLGRSDGAVTQAPPPAPMAARRSQRGRGEGTALLVPLA 60
 LALACLGLLLAVSLGSRASLSAQEELVAEEDQDPSELNPQTESQDPAPFL 120
 LALACLGLLLAVSLGSRASLSAQEELVAEEDQDPSELNPQTESQDPAPFL 120
 RPRRSAPKGRKTRARRATAAHYEVHPRPGDGAQAGVDGTVSGWEEARINSSPLR 180
 RPRRSAPKGRKTRARRATAAHYEVHPRPGDGAQAGVDGTVSGWEEARINSSPLR 180
 IGEFIVTRAGLYLYLCVHFDEGKAVYIKDLLVDGVLALRCLEEFSAASSLGP 240
 IGEFIVTRAGLYLYLCVHFDEGKAVYIKDLLVDGVLALRCLEEFSAASSLGP 240
 CQVSGLLALRPGSSLRINTLPWAHLKAAPFLTYFGLFQVH 284
 CQVSGLLALRPGSSLRINTLPWAHLKAAPFLTYFGLFQVH 284

dard; protein; 249 AA.

(first entry)

necrosis factor Apo-3 ligand protein sequence.

necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
 ndent transcription; JNK/SAPK-dependent response; cancer.

98WC-US021407.

97US-0062037P.

97US-0069862P.

TECH INC.

Marsters SA, Pitti R;

1982/24.

5000.

3- ligand (a tumor necrosis factor) homologue.

1; 74pp; English.

sequence represents a human tumour necrosis factor (TNF) and

CC Lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcripti
 CC to induce JNK/SAPK-dependent responses in mammalian cells
 XX
 SQ Sequence 249 AA;

Query Match 87.8%; Score 1268; DB 2; Length 249;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; G

Qy 36 MAARRSORRRGRGEGTALLVPLALGLGLALACLGALLAVSLGSRASLSAQEPA
 Db 1 MAARRSORRRGRGEGTALLVPLALGLGLALACLGALLAVSLGSRASLSAQEPA
 Qy 96 VAEDQDPSELNPQTESQDPAPFLNLRVPRRSAPKGRKTRARRATAAHYEVHPR
 Db 61 VAEDQDPSELNPQTESQDPAPFLNLRVPRRSAPKGRKTRARRATAAHYEVHPR
 Qy 156 GAQAGVDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYLCVHFDEGKAVY
 Db 121 GAQAGVDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYLCVHFDEGKAVY
 Qy 216 LLYDGVIALRCLEEFSAASSLGPQLRCQVSGLLALRPGSSLRINTLPWAHLKA
 Db 181 LLYDGVIALRCLEEFSAASSLGPQLRCQVSGLLALRPGSSLRINTLPWAHLKA
 Qy 276 TYFGLFQVH 284
 Db 241 TYFGLFQVH 249

RESULT 3

AAAY95338

ID AAAY95338 standard; protein; 249 AA.

AC AAAY95338;

XX 25-SEP-2000 (first entry)

XX Human PRO207 antitumour protein.

XX PRO207; human; antitumour; tumour; therapy; cytostatic; breast ca
 KW ovarian cancer; renal cancer; colorectal cancer; uterine cancer;
 KW prostate cancer; lung cancer; bladder cancer;
 KW central nervous system cancer; melanoma; leukaemia; neoplasia.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..40

FT /label= Signal_peptide

FT Modified-site 10..14

FT /note= "amidation"

FT Peptide 24..35

FT /note= "prokaryotic membrane lipoprotein lipid"

FT Modified-site 27..33

FT /note= "N-myristoylation"

FT Modified-site 29..35

FT /note= "N-myristoylation"

FT Modified-site 36..42

FT /note= "N-myristoylation"

FT Protein 41..249

FT /label= PRO207

FT Modified-site 45..51

FT /note= "N-myristoylation"

FT Modified-site 97..101

FT /note= "amidation"

FT Modified-site 118..124

FT /note= "N-myristoylation"

FT Modified-site 121..127

FT /note= "N-myristoylation"

FT Modified-site 125..131

/note= "N-myristoylation"
128. .134
/note= "N-myristoylation"
139. .143
/note= "Asn is N-glycosylated"

99WO-US028565.

98US-0113296P.

99WO-US005028.

98US-0130232P.

98US-0131443P.

98US-0134287P.

98US-0144758P.

98US-0145698P.

99WO-US021090.

99WO-US021547.

ECH INC.

Goddard A., Godowski P., Gurney AL., Marsters SA;

tti RM, Wood WI;

58/38.

17.

ion to inhibit neoplastic cell growth or for treating tumor
ries polypeptides PRO179, PRO207, PRO320, PRO219, PRO221,
, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866.

4; 172pp; English.

quence is that of human antitumor protein PRO207, as
foetal kidney cDNA clone (see AAA49717). PRO207 shows
ence identity to tumour necrosis factor family members,
an lymphotoxin-beta (23.4%) and human CD40 ligand (19.8%).
216. A claimed method for inhibiting the growth of a tumour
exposing the tumor cell to PRO179, PRO207, PRO320, PRO219,
, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866
49), their agonists or chimeric polypeptides incorporating
ur is especially a cancer selected from breast, ovarian,
tal, uterine, prostate, lung, bladder and central nervous
melanoma and leukaemia. Methods for the recombinant
the antitumor proteins are also provided

A;

87.8%; Score 1268; DB 3; Length 249;
rity 100.0%; Pred. No. 1.3e-112; Indels 0; Gaps 0;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
SQRRRGRGPGTALLVPLALGLALACGLLLAVVSLGSRASLSAQEEL 95
SQRRRGRGPGTALLVPLALGLALACGLLLAVVSLGSRASLSAQEEL 60
QDPSELNPQTEESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYVHP 155
QDPSELNPQTEESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYVHP 120
VDGTVSGWEEARINSSPLRYNQIGEFIVTRAGLYLYCOVHDEGKAVYLKLD 215
VDGTVSGWEEARINSSPLRYNQIGEFIVTRAGLYLYCOVHDEGKAVYLKLD 180
VLAIRCLFEFSATASLGPQLRCQVSGLLALPGSSLRIRTLTPWAHLKAPFL 275
VLAIRCLFEFSATASLGPQLRCQVSGLLALPGSSLRIRTLTPWAHLKAPFL 240
FQVH 284
|||

Db 241 TYFGLFQVH 249

RESULT 4
AAB07526
ID AAB07526 standard; protein; 249 AA.
XX AAB07526;
AC AAB07526;
XX 20-OCT-2000 (first entry)
XX 20-OCT-2000 (first entry)
XX Amino acid sequence of a soluble recombinant human TWEAK protein.
DE TWEAK protein; immunological disorder; immune response; inflammati
KW TWEAK blocking agent; autoimmune disease; organ transplant rejecti
KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock;
XX Homo sapiens.
OS WO200042073-A1.
XX 20-JUL-2000.
PD 20-JUL-2000.
XX 14-JAN-2000; 2000WO-US001044.
XX 15-JAN-1999; 99US-0116168P.
XX (BIOJ) BIOGEN INC.
PA Renmert P;
PI WPI; 2000-476036/41.
XX Preventing and treating immune responses using modulators, especia
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful fo
PT treating e.g. inflammation and graft versus host disease.
XX Disclosure; Fig 1; 45pp; English.
XX The present sequence represents a TWEAK protein. The specification
describes a method for preventing or treating an immunological dis
and/or inhibiting an immune response in an animal. The method comp
administering a TWEAK blocking agent. The method may be used for
preventing and treating immune disorders associated with inappropri
expression and/or activity of TWEAK. These disorders include autoi
diseases, acute and chronic inflammation, organ transplant rejecti
Graft-versus-Host disease (GVHD), lymphoid cell malignancies, sept
other forms of shock, loss of immune responsiveness (as seen in hu
immunodeficiency virus (HIV) infections) and failure of the immune
response to tumour growth
XX Sequence 249 AA;
Query Match 87.8%; Score 1268; DB 3; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.3e-112; Indels 0; G
Matches 249; Conservative 0; Mismatches 0; Indels 0; G
QY 36 MAARSQRRGRGPGTALLVPLALGLALACGLLLAVVSLGSRASLSAQEPAQ
Db 1 MAARSQRRGRGPGTALLVPLALGLALACGLLLAVVSLGSRASLSAQEPAQ
QY 96 VAEEDQDPSELNPQTEESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYVHPRE
Db 61 VAEEDQDPSELNPQTEESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYVHPRE
QY 156 GAQAGVDGTVSGWEEARINSSPLRYNQIGEFIVTRAGLYLYCOVHDEGKAVYI
Db 121 GAQAGVDGTVSGWEEARINSSPLRYNQIGEFIVTRAGLYLYCOVHDEGKAVYI
QY 216 LIAVDGVLALRCLFEFSATASLGPQLRCQVSGLLALPGSSLRIRTLTPWAHLKAF
Db 181 LIAVDGVLALRCLFEFSATASLGPQLRCQVSGLLALPGSSLRIRTLTPWAHLKAF

FOVH 284
|||||
FOVH 249

iard; protein; 249 AA.

(first entry)
polypeptide.

enign tumour; malignant tumour; lymphoid malignancy;
uronal disorder; stromal disorder; blastocoealic disorder;
disorder; immune disorder; angiogenic disorder; cytostatic;
ve.

1.

2000WO-US003565.

99WO-US005028.
99US-0123972P.
99US-0133459P.
99WO-US012252.
99US-0140650P.
99US-0140653P.
99US-0144758P.
99US-0145698P.
99US-0146222P.
99US-0149395P.
99US-0151689P.
99WO-US020111.
99WO-US021090.
99WO-US028313.
99WO-US028301.
99WO-US028634.
2000WO-US000219.

TECH INC.

Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
Wood WI;

567/26.
255.

nucleic acids encoding PRO polypeptides, useful for treating
ignant tumors, leukemias and lymphoid malignancies,
angiogenic and immunologic disorders.

4; 302pp; English.

vention relates to the isolation of novel human PRO
and the polynucleotide sequences encoding them. The PRO
agonists, antagonists or anti-PRO antibodies are useful for
gn or malignant tumours (e.g. renal, kidney, bladder, such
leukaemias and lymphoid malignancies, other disorders such
glial, astrocytal, hypothalamic, glandular, macrophagal,
blastocoealic disorders, inflammatory, immune and angiogenic
e polynucleotide sequences are also useful in gene therapy.
16162 represent the human PRO polypeptides of the invention

AA;

Query Match 87.8%; Score 1268; DB 5; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 249; Conservative 0; Mismatches 0; Indels 0; G

QY 36 MAARRSQRGRGEGPTALLVPLALGLALACLGILLAVSLGSRASLSAQBPA
Db 1 MAARRSQRGRGEGPTALLVPLALGLALACLGILLAVSLGSRASLSAQBPA
QY 96 VAEEDQDPSLNPQTEESQDPAPFLNRLVRPRSPKGRKTRARRAIAAHYEVHPR
Db 61 VAEEDQDPSLNPQTEESQDPAPFLNRLVRPRSPKGRKTRARRAIAAHYEVHPR
QY 156 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYLCQVHFDGKAVY
Db 121 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYLCQVHFDGKAVY
QY 216 LLDVGVLAALCLFEFSATAASSLGPQLRLCQVSGLLALRPQSSLRITLPAHLKA
Db 181 LLDVGVLAALCLFEFSATAASSLGPQLRLCQVSGLLALRPQSSLRITLPAHLKA
QY 276 TYFGLFQVH 284
Db 241 TYFGLFQVH 249

RESULT 6

ABR42315
ID ABR42315 standard; protein; 249 AA.

XX AC ABR42315;

DT 11-AUG-2003 (first entry)

XX DE Human TWEAK protein.

XX KW Human; TWEAK; tumour necrosis factor; ligand; cytostatic;
XX OS immunomodulator; osteopathic.

XX OS Homo sapiens.

PN WO2003040307-A2.

XX PD 15-MAY-2003.

XX PF 25-JUL-2002; 2002WO-US023782.

XX PR 27-JUL-2001; 2001US-0307838P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PT Hilbert DH, Rosen CA;

XX DR WPI; 2003-430659/40.

XX DR N-PSDB; ACC57901.

XX PT New heteromultimeric complex having a first polypeptide member of
tumor necrosis factor (TNF) ligand family, and a second different
of TNF ligand family, useful for treating cancer, osteoporosis or
autoimmune disease.

XX PS Disclosure; Page 368-369; 388pp; English.

XX CC The present sequence is the protein sequence for human TWEAK prot
invention relates to compositions comprising heterotrimeric compl
tumor necrosis factor (TNF) ligand family members, and their use
detection, prevention and treatment of disease. In one embodiment
heterotrimeric complex comprises full-length or extracellular por
TWEAK and full-length or extracellular portions of other TNF liga
family members, preferably VEGI or VEGI-SV. The heterotrimeric co
of the invention are useful for treating an autoimmune disease, c
osteoporosis, and particularly for inhibiting cancer cell prolif
increasing B cell proliferation, or inducing apoptosis of T cells

XX

87.8%; Score 1268; DB 6; Length 249;
 city 100.0%; Pred. No. 1.3e-112; Indels 0; Gaps 0;
 iservative 0; Mismatches 0; Indels 0; Gaps 0;
 SQRGRGRGPGTALLVPLALGLGLALACGLLLAVVSLGSRASLSAQEPQEL 95
 SQRGRGRGPGTALLVPLALGLGLALACGLLLAVVSLGSRASLSAQEPQEL 60
 QDSELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRQD 155
 QDSELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRQD 120
 VDGTVSGWEERINSSPLRYNROI GEFIVTRAGLYLYLCQVHDEGKAVYLKLD 215
 VDGTVSGWEERINSSPLRYNROI GEFIVTRAGLYLYLCQVHDEGKAVYLKLD 180
 VLALRCLEEPSATAASLGQRLCQVSGLLALRPGSSLRIRTLPMWHLKAAPPL 275
 VLALRCLEEPSATAASLGQRLCQVSGLLALRPGSSLRIRTLPMWHLKAAPPL 240
 FQVH 284
 FQVH 249

ard; protein; 249 AA.

first entry)

nd family member #12.

necrosis factor; TNF ligand; endokine alpha;
 resorption disorder; osteoporosis; Paget's disease;
 fication.

1.

002US-00218547.

001US-0312542P.

001US-0330761P.

C A.

LI B.

Rosen CA, Nardelli B;

172/66.

105.

alpha gene useful for preparing a composition for treating a
 iated with excessive or insufficient bone resorption e.g.,
 Paget's disease or arterial calcification.

3Q ID NO 24; 145pp; English.

relates to an isolated nucleic acid molecule encoding a
 is factor family ligand. A composition comprising the
 body or its fragment is used for treating an individual in
 ased level of endokine alpha activity. The endokine alpha
 resent in a heterotrimeric complex is used for treating an

CC individual having a disorder associated with excessive bone resorp
 e.g. osteoporosis, Paget's disease or arterial calcification. Trea
 CC individual having a disorder associated with insufficient bone res
 CC comprises administering an endokine alpha antagonist, which is the
 CC antibody that binds specifically to endokine alpha polypeptide. Th
 CC present sequence represents the amino acid sequence of a tumour ne
 CC factor family ligand.

SQ Sequence 249 AA;

Query Match 87.8%; Score 1268; DB 7; Length 249;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 36 MAARRSQRRGRGPGTALLVPLALGLGLALACGLLLAVVSLGSRASLSAQEPAC
 DB 1 MAARRSQRRGRGPGTALLVPLALGLGLALACGLLLAVVSLGSRASLSAQEPAC
 QY 96 VAERDQDSELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRF
 DB 61 VAERDQDSELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRF
 QY 156 GQAQGVDTGTVSGWEERINSSPLRYNROI GEFIVTRAGLYLYLCQVHDEGKAVYI
 DB 121 GQAQGVDTGTVSGWEERINSSPLRYNROI GEFIVTRAGLYLYLCQVHDEGKAVYI
 QY 216 LLVDGVLALRCLEEPSATAASLGQRLCQVSGLLALRPGSSLRIRTLPMWHLKAF
 DB 181 LLVDGVLALRCLEEPSATAASLGQRLCQVSGLLALRPGSSLRIRTLPMWHLKAF
 QY 276 TYFGLFQVH 284
 DB 241 TYFGLFQVH 249

RESULT 8

AAW29745

ID AAW29745 standard; protein; 249 AA.

XX AC AAW29745;

DT 27-OCT-1998 (first entry)

DE TNF related endothelium proliferative agent protein.

XX TNF; endothelium proliferative agent; TREPA; wound healing; cance.
 XX tissue grafting; vascularisation; apoptosis; autoimmune; birth co

OS HOMO sapiens.

XX WO9835061-A2.

XX 13-AUG-1998.

XX 12-FEB-1998; 98WO-US002859.

XX 12-FEB-1997; 97US-00798692.

XX 10-FEB-1998; 98US-00021706.

PA (ABBO) ABBOTT LAB.

XX Wiley SR;

XX WPI; 1998-447255/38.

DR N-PSDB; AAV47613.

XX Detecting nucleic acid encoding TREPA - useful for diagnosis and
 PT treatment of autoimmune disease, tumours and inflammation.

XX Claim 16; Page 123-4; 142pp; English.

XX The TNF-related endothelium proliferative agent (TREPA), or its
 CC activators or agonists, are used to treat a deficit of TREPA, e.g

healing or tissue grafting, by promoting vascularisation, apoptosis for treating cancer and eliminating autoreactive cells adjunct to cancer chemotherapy or antiviral treatment. A can also be used to target cytotoxic agents or for action of the corresponding receptor, the nucleic acid for used to transform tumour cells to render them more TREPA and to screen for TREPA mimics. Ribozymes, antisense ies or peptides, are used to treat TREPA-associated . tumours and metastases (by inhibiting vascularisation), or a wide range of autoimmune conditions, conditions ormal stimulation of epithelial cells (e.g. is), for birth control (inhibiting ovulation and placental other angiogenic conditions (e.g. ulcers)

AA;

87.6%; Score 1265; DB 2; Length 249;

arity 99.6%; Pred. No. 2.5e-112;

onservative 1; Mismatches 0; Indels 0; Gaps 0;

RSQRRGRGPGCTALLVPLALGGLALACGLLLAVVSLGSRASLSAQBPQEL 95

RSQRRGRGPGCTALLVPLALGGLALACGLLLAVVSLGSRASLSAQBPQEL 60

DQDPSLNPOEESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHVEVHPRPQD 155

DQDPSLNPOEESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHVEVHPRPQD 120

GVDTGTVSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAVYLKLD 215

GVDTGTVSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAVYLKLD 180

GVDTGTVSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAVYLKLD 275

GVDTGTVSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAVYLKLD 240

LFQVH 284

LFQVH 249

(first entry)

(TNF related endothelium proliferative agent).

necrosis factor; TNF; angiogenesis; wound healing; TREPA;

endothelium proliferative agent; tumour; metastasis;

Location/Qualifiers

98..249

/label= Extracellular_domain

98US-00105343.

97US-00798692.

98US-00021706.

TT LAB.

XX WPI; 2001-280760/29.
DR N-PSDB; AAD04350.
XX Inducing angiogenesis in mammal at desired sites for promoting wo
PT healing, by administering soluble fragment of extracellular domai
PT tumor necrosis factor related endothelium proliferative agent pro
XX Claim 1; Col 75-76; 53pp; English.

XX The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molec
CC designated as TREPA (TNF related endothelium proliferative agent)
CC Soluble biologically active TREPA are used to treat TREPA-associ
CC diseases, tumours or metastases. TREPA is used for inducing angic
CC in human for promoting wound healing and for vascularising grafted
CC for successful grafting and to promote tissue grafts. The present
CC acid sequence is clone ID #690050 human TREPA

SQ Sequence 249 AA;

Query Match 87.6%; Score 1265; DB 4; Length 249;

Best Local Similarity 99.6%; Pred. No. 2.5e-112;

Matches 248; Conservative 1; Mismatches 0; Indels 0; G

QY 36 MAARRSQRGRGPGCTALLVPLALGGLALACGLLLAVVSLGSRASLSAQBPQ

DB 1 MAARRSQRGRGPGCTALLVPLALGGLALACGLLLAVVSLGSRASLSAQBPQ

QY 96 VAEDQDPSELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHVEVHP

DB 61 VAEDQDPSELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAIAAHVEVHP

QY 156 QAQAGVDTGTVSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAV

DB 121 QAQAGVDTGTVSGWEARINSSPLRYNQIGEFIVTRAGLYLYCQVHFDEGKAV

QY 216 LLVDGVLAALRCLEEFSAATAASSLGPQLRQCQVSGLLALRPGSSLRIRTLPAHLK

DB 181 LLVDGVLAALRCLEEFSAATAASSLGPQLRQCQVSGLLALRPGSSLRIRTLPAHLK

QY 276 TYFGLFQVH 284

DB 241 TYFGLFQVH 249

RESULT 10

ADC97712

ID ADC97712 standard; protein; 249 AA.

XX ADC97712;

XX 15-JAN-2004 (first entry)

XX Murine FL-TWEAK.

XX Murine; FL-TWEAK; TNF relatedness and weak ability to induce cel.

XX TNF; Tumour Necrosis Factor; TWEAK; fibrosis; cardiac disease;

XX liver disease; lung disease; kidney disease; skin disease;

XX skeletal muscle disease; adipose tissue disease;

XX gastrointestinal tract disease; pancreatic disease;

XX reproductive organ disease; neural disease; cartilage disease;

XX bone disease; connective tissue disease; cellular death; hepatot;

XX dermatological; gastrointestinal; osteopathic.

XX Mus sp.

XX WO2003086311-A2.

XX 23-OCT-2003.

XX 09-APR-2003; 2003WO-US011350.

XX

hard; protein; 208 AA.

(first entry)

rotein.

is factor receptor; signal transducer molecule; TNF; APO4; abnormality; gestational abnormality; prostate cancer; PO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease; main; immunogen; antibody preparation; breast carcinoma; man.

98WO-US018393.

97US-00924634.

WASHINGTON.

1191/17.

1424.

rosis Factor family receptor polypeptides and ligands - agnosis and treatment of prostate cancer and developmental abnormalities.

13A; 156pp; English.

on describes isolated Tumor Necrosis Factor (TNF) family /peptides: APO4, APO6, APO8 and APO9 or their active id isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or fragments. APO4 is useful for diagnosing prostate cancer by levels of APO4 in an individual. Prostate cancer can also be 3 APO4 selective binding agents linked to a therapeutic polypeptides are also useful for identifying selective ts useful in diagnosis/treatment of disease by binding of a polypeptide/active fragment which is extracellular. or the cell surface. The binding is preferably performed in clypeptides/ active fragments are also useful for screening and antagonists by binding and observing the changer in APO4 fective pharmacological agents useful in diagnosis or disease are also identified using APO4 polypeptides/active d APO4 signal transducer molecules that specifically interact laemic domain of APO4 and detecting a change in level of APO4 e method is performed in vivo or in vitro. APO polypeptides ul as immunogens for preparing antibodies. APO4 is also agnosis/treatment of developmental or gestational s. APO8 was transfected to human breast carcinoma cell line induced apoptosis

AA;

73.5%; Score 1062; DB 2; Length 208;

larity 99.5%; Pred. No. 4.9e-93;

Conservative 0; Mismatches 1; Indels 0; Gaps 0;

GSRLSQAQAEELVAEEQDPSELNPQTEESQDPAPFLNLRVRRSAPKGRKT 136

GSRLSQAQAEELVAEEQDPSELNPQTEESQDPAPFLNLRVRRSAPKGRKT 60

RAIAAHYVHRPQDGAQAGVDGTVSGWEARINSSPLRYNQIGEFIVTRAGLY 196

RAIAAHYVHRPQDGAQAGVDGTVSGWEARINSSPLRYNQIGEFIVTRAGLY 120

197 YLYCQVHFDEGKAVYKLDLLVDGVLALRCLEEFSAATAASLGQRLCQVSGLLA

121 YLYCQVHFDEGKAVYKLDLLVDGVLALRCLEEFSAATAASLGQRLCQVSGLLA

257 SSLRIRTLPAHLKAAPFLTYEGLFQVH 284

181 SSLRIRTLPAHLKAAPFLTYEGLFQVH 208

RESULT 13

AAW47524

ID AAW47524 standard; protein; 225 AA.

XX AC AAW47524;

XX 21-JUL-1998 (first entry)

XX Mus musculus tumour necrosis factor related ligand (TRELL).

XX TRELL; tumour necrosis factor related ligand; tnfr; treatment; car

XX auto-immune disease; immune system; stimulation; suppression;

XX graft rejection.

XX Mus musculus.

XX Key Location/Qualifiers

XX FT Domain 1..21

XX /note= "hydrophobic, transmembrane domain"

XX WO9805783-A1.

XX 12-FEB-1998.

XX 07-AUG-1997; 97WO-US013945.

XX 07-AUG-1996; 96US-0023541P.

XX 18-OCT-1996; 96US-0028515P.

XX 18-MAR-1997; 97US-0040820P.

XX (BIOJ) BIOGEN INC.

XX (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX Chicheportiche Y, Browning JL;

XX WPI; 1998-145619/13.

XX N-PSDB; AAV18599.

XX Tumour necrosis factor related ligand - useful for, e.g. treatin

XX auto-immune disease and immune responses to tissue grafts.

XX Claim 12; Page 48-50; 69pp; English.

XX The sequence is that of mouse tumour necrosis factor related lig

XX (TRELL). TRELL or active fragments can be included with a carrie

XX pharmaceutical compositions to treat cancer, autoimmune diseases

XX immune responses to tissue grafts, or to stimulate or suppress t

XX system. It is useful to screen for TRELL receptors, by labelling

XX detectable label and screening compositions for binding. Agents

XX interfering with TRELL-receptor binding can also be screened for

XX then be administered, optionally with interferon- gamma, to indu

XX death or treat, suppress or alter immune responses (especially i

XX human adenocarcinoma cells) involving a signal pathway between t

XX its receptor. It's coding sequence can be used in gene therapy f

XX related disorders in mammals (especially humans), e.g. tumours,

XX autoimmune and inflammatory diseases or inherited genetic disord

XX introducing into cells, and expressing, therapeutically effectiv

XX of a vector, e.g. a virus comprising a gene encoding TRELL. It n

XX be of use in the preparation of prepare probes for screening

XX natural/synthetic DNAs for TRELL-encoding sequences and for anti

XX therapy

XX Sequence 225 AA;

70.6%; Score 1020; DB 2; Length 225;
 88.8%; Pred. No. 5.7e-89;
 9; Mismatches 16; Indels 0; Gaps 0;
 ALACGLLLAVVSLGSRASLSAQEPAPQBELVAEEDQDPSELNPQTESQDPE
 ALACGLLLVVVSLGSMATLSAQEPSELTAEDRREPPELNPQTESQDVVPEL 61
 PRSAPKGRKTRARAAAHYEVHPRPGDGAQAGVDGTVSGWEARINSSPLR 180
 PRSAPKGRKTRARAAAHYEVHPRPGDGAQAGVDGTVSGWEARINSSPLR 121
 GEFTVIRAGLYLYCQVHDEGKAVYLLKDLVGVLAALRCLEEFSAATAASLGP 240
 GEFTVIRAGLYLYCQVHDEGKAVYLLKDLVGVLAALRCLEEFSAATAASLGP 181
 QVSGLLALRPGSSLRIRTPWAHLKAAPFLTYFGLFQVH 284
 QVSGLLALRPGSSLRIRTPWAHLKAAPFLTYFGLFQVH 225

iard; protein; 225 AA.

(first entry)

quence of a soluble recombinant murine TWEAK protein.

; immunological disorder; immune response; inflammation;
 3 agent; autoimmune disease; organ transplant rejection;
 host disease; GVHD; lymphoid cell malignancy; shock; tumour.

1.

2000WO-US001044.

99US-0116168P.

N INC.

036/41.

d treating immune responses using modulators, especially
 f TWEAK, TWEAK receptors and TWEAK ligands, useful for
 inflammation and graft versus host disease.

ig 1; 45pp; English.

sequence represents a TWEAK protein. The specification
 method for preventing or treating an immunological disorder
 ting an immune response in an animal. The method comprises
 a TWEAK blocking agent. The method may be used for
 id treating immune disorders associated with inappropriate
 id/or activity of TWEAK. These disorders include autoimmune
 re and chronic inflammation, organ transplant rejection,
 Host disease (GVHD), lymphoid cell malignancies, septic and
 f shock, loss of immune responsiveness (as seen in human
 ency virus (HIV) infections) and failure of the immune
 tumour growth

AA;

70.6%; Score 1020; DB 3; Length 225;

Best Local Similarity 88.8%; Pred. No. 5.7e-89;
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;
 QY 61 LGIGLALACGLLLAVVSLGSRASLSAQEPAPQBELVAEEDQDPSELNPQTESQDPE
 DB 2 LSIGLALACGLLLVVVSLGSMATLSAQEPSELTAEDRREPPELNPQTESQDVV
 QY 121 NLRVPRRSPKGRKTRARAAAHYEVHPRPGDGAQAGVDGTVSGWEARINSS
 DB 62 EQLVPRRSPKGRKTRARAAAHYEVHPRPGDGAQAGVDGTVSGWEARINSS
 QY 181 YRQIGFEFTVIRAGLYLYCQVHDEGKAVYLLKDLVGVLAALRCLEEFSAATAAS
 DB 122 YRQIGFEFTVIRAGLYLYCQVHDEGKAVYLLKDLVGVLAALRCLEEFSAATAAS
 QY 241 QRLCQVSGLLALRPGSSLRIRTPWAHLKAAPFLTYFGLFQVH 284
 DB 182 QRLCQVSGLLALRPGSSLRIRTPWAHLKAAPFLTYFGLFQVH 225

RESULT 15

AAW93591

ID AAW93591 standard; protein; 211 AA.

XX AC AAW93591;

XX DT 18-JUN-1999 (first entry)

XX DE Mouse TNRL3 protein.

XX KW Tumour necrosis factor receptor; signal transducer molecule; TNF;
 KW developmental abnormality; gestational abnormality; prostate ca
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carci
 KW apoptosis; mouse.

XX OS Mus sp.

XX PN WO9911791-A2.

XX PD 11-MAR-1999.

XX PF 04-SEP-1998; 98WO-US018393.

XX PR 05-SEP-1997; 97US-00924634.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Chaudhary PM;

XX DR WPI; 1999-205191/17.

XX DR N-PSDB; AAX23425.

XX PT New Tumor Necrosis Factor family receptor polypeptides and ligand
 PT useful for diagnosis and treatment of prostate cancer and develop
 PT or gestational abnormalities.

XX PS Claim 40; Fig 13B; 156pp; English.

XX CC This invention describes isolated Tumor Necrosis Factor (TNF) fan
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TN
 CC their active fragments. APO4 is useful for diagnosing prostate ca
 CC determining levels of APO4 in an individual. Prostate cancer can
 CC treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selecti
 CC binding agents, useful in diagnosis/treatment of disease by bindi
 CC agents to the polypeptide/active fragment which is extracellular,
 CC expressed on the cell surface. The binding is preferably perform
 CC vivo. APO4 polypeptides/ active fragments are also useful for sc
 CC for agonists and antagonists by binding and observing the change
 CC activity. Effective pharmacological agents useful in diagnosis o
 CC treatment of disease are also identified using APO4 polypeptides,

06:25:22 2004

us-09-245-198a-4.rag

AP04 signal transducer molecules that specifically interact with the cytoplasmic domain of AP04 and detecting a change in level of AP04. The method is performed *in vivo* or *in vitro*. AP04 polypeptides are used as immunogens for preparing antibodies. AP04 is also used as a target for diagnosis, prognosis/treatment of developmental or gestational disorders. AP08 was transfected to human breast carcinoma cell line MCF-7. AP08 was used to study the role of AP08 in cell induced apoptosis.

AA;

```

67.0%; Score 968; DB 2; Length 211;
arity 89.1%; Pred. No. 4.8e-84;
conservative 9; Mismatches 14; Indels 0; Gaps 0;

```

```

|SLGSRASLSAQEPAQEELVAEEDQDPSELNPQTEESQDPAPFLNLRVPRRSAPKG 133
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
|SLGSWATLSAQEPSOEELTAEDRREPELNPQTEESQDVVFLEQLVRPRRSAPKG 60

```

PARRAIAAHYEVHPRPGDGAQAGVDGTVSGWEEARINSSSPLYRNQIGEFIVTRA 193
 :
 PRRAIAAHYEVHPRPGDGAQAGVDGTVSGWEEETKINS SSPLYDRQIGEFTVIRA 120

LYCQVHDEGKAVYLKDLLVDGVIALRCLEEFSAATAASSLGPQIRLCQVSGLLAL 253

LYCQVHDEGKAVYLKDLLVNGVIALRCLEEFSAATAASSPGQIRLCQVSGLLAL 180

SSLRIRTL PWAHLKAAPFLTYFGLFQVH 284
SSLRIRTL PWAHLKAAPFLTYFGLFQVH 211

April 7, 2004, 17:44:47
5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

LC search, using sw model

til 8, 2004, 19:08:51 : Search time 5584 Seconds
(without alignments)
10657.232 Million cell updates/sec

-09-245-198A-3

atgtcattgttagacttga.....gacaaaattgtgataaatgg 1373

IGO_NUC

pop 60.0 , Gapext 60.0

70272 seqs, 21671516995 residues

ts satisfying chosen parameters: 6940544

gth: 0

gth: 2000000000

listing first 100 summaries

enEmbl:

gb_ba:

gb_htg:

gb_in:

gb_om:

gb_ov:

gb_pat:

gb_ph:

gb_pl:

gb_ro:

gb_sts:

gb_sy:

gb_un:

gb_vi:

em_ba:

em_fun:

em_hum:

em_in:

em_mu:

em_or:

em_ov:

em_pat:

em_ph:

em_pl:

em_ro:

em_sts:

em_un:

em_vi:

em_htg_hum:

em_htg_inv:

em_htg_other:

em_htg_mus:

em_htg_pln:

em_htg_rod:

em_htg_mam:

em_htg_vrt:

em_sy:

em_htgo_hum:

em_htgo_mus:

em_htgo_other:

; the number of results predicted by chance to have a

score greater than or equal to the score of the result being pri
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1373	100.0	1373	6	BD062758	BD062758 A
2	1285	93.6	1306	9	AF030099	AF030099 Hoi
3	1172	85.4	1353	6	AX201324	AX201324 Se
4	1172	85.4	1353	9	AY358870	AY358870 Hoi
5	1172	85.4	1368	9	AF055872	AF055872 Hoi
6	1172	85.4	1421	6	BD090952	BD090952 Ap
7	958	69.8	1236	6	AR140407	AR140407 Se
8	958	69.8	1236	6	BD057124	BD057124 Me
9	796	58.0	1642	6	BC019047	BC019047 Hoi
10	625	45.5	898	6	AX180714	AX180714 Se
11	620	45.2	60268	9	AC016876	AC016876 Hoi
12	553	40.3	1816	9	AY081051	AY081051 Hoi
13	435	31.7	218485	2	AC127470	AC127470 Pa
14	179	13.0	195	6	AX379024	AX379024 Se
15	145	10.6	180222	2	AC130192	AC130192 Su
16	119	8.7	130254	2	AC136195	AC136195 Ra
17	119	8.7	165316	2	AC119115	AC119115 Ra
18	119	8.7	203083	2	AC069459	AC069459 Mu
19	119	8.7	203083	2	AC098923	AC098923 Ra
20	119	8.7	203083	2	AC136563	AC136563 Ra
21	119	8.7	203083	2	AL603707	AL603707 M
22	119	8.7	203083	2	AC126237	AC126237 Ca
23	119	8.7	203083	2	AF030100	AF030100 M
24	119	8.7	203083	2	AC126921	AC126921 Bo
25	119	8.7	203083	2	AC128925	AC128925 Ca
26	119	8.7	203083	2	AC126239	AC126239 Fe
27	119	8.7	203083	2	AX201395	AX201395 Se
28	119	8.7	203083	2	BD090954	BD090954 Af
29	119	8.7	203083	2	AC129071	AC129071 Pa
30	119	8.7	203083	2	BD062757	BD062757 A
31	119	8.7	203083	2	AC118309	AC118309 Ra
32	119	8.7	203083	2	AX201459	AX201459 Se
33	119	8.7	203083	2	AX522872	AX522872 Se
34	119	8.7	203083	2	AX522345	AX522345 Se
35	119	8.7	203083	2	AX522465	AX522465 Se
36	119	8.7	203083	2	AX322295	AX322295 Se
37	119	8.7	203083	2	AX209050	AX209050 Se
38	119	8.7	203083	2	BC062447	BC062447 Hc
39	119	8.7	203083	2	AR427020	AR427020 Se
40	119	8.7	203083	2	BD122573	BD122573 Ec
41	119	8.7	203083	2	AX198506	AX198506 Se
42	119	8.7	203083	2	AR427022	AR427022 Se
43	119	8.7	203083	2	BD122575	BD122575 Ec
44	119	8.7	203083	2	BC032480	BC032480 Hc
45	119	8.7	203083	2	AX330518	AX330518 Se
46	119	8.7	203083	2	AX410097	AX410097 Se
47	119	8.7	203083	2	AX895134	AX895134 Se
48	119	8.7	203083	2	BD030667	BD030667 Se
49	119	8.7	203083	2	AX381620	AX381620 Se
50	119	8.7	203083	2	BD224095	BD224095 Se
51	119	8.7	203083	2	AX408082	AX408082 Se
52	119	8.7	203083	2	BC060315	BC060315 Hc
53	119	8.7	203083	2	AR413525	AR413525 Se
54	119	8.7	203083	2	BD109078	BD109078 Ec
55	119	8.7	203083	2	AX885407	AX885407 Se
56	119	8.7	203083	2	BD025017	BD025017 Se
57	119	8.7	203083	2	AR415934	AR415934 Se
58	119	8.7	203083	2	BD111487	BD111487 Ec
59	119	8.7	203083	2	AX381939	AX381939 Se
60	119	8.7	203083	2	BC062222	BC062222 Hc
61	119	8.7	203083	2	AR415931	AR415931 Se
62	119	8.7	203083	2	BD111484	BD111484 Ec
63	119	8.7	203083	2	AX885406	AX885406 Se
64	119	8.7	203083	2	BD025016	BD025016 Se
65	119	8.7	203083	2	BD224094	BD224094 Se

1.9 555 6 AR415935 Sequence
1.9 555 6 BD111488 EST and e
1.9 560 9 BC042807 Homo sapi
1.9 595 6 AR415932 Sequence
1.9 595 6 BD111485 EST and e
1.9 1822 9 AK130060 Homo sapi
1.9 2183 9 HS802687 Homo sapi
1.9 2650 9 AK124175 Homo sapi
1.9 110293 9 HSDJ686N3 Human DNA
1.9 159020 2 AC027023 Homo sapi
1.8 507 6 AR427029 Sequence
1.7 40 6 BD122582 EST and e
1.7 40 6 BD090955 Apo-3 lig
1.7 478 6 AR427041 Sequence
1.7 478 6 BD122594 EST and e
1.7 495 6 AR427028 Sequence
1.7 495 6 BD122581 EST and e
1.7 23 6 AX201458 Sequence
1.7 164117 5 AL845513 Zebrafish
1.7 172618 2 AC144822 Danio rer
1.7 214279 2 AC138245 Mus muscu
1.7 248178 2 AC093996 Rattus no
1.6 38 6 BD090956 Apo-3 lig
1.6 1005 6 AX695886 Sequence
1.6 1173 1 AF182516 Pseudomon
1.6 2942 6 AX695885 Sequence
1.6 2942 10 MWU250723 Mus muscu
1.6 3642 10 BC046524 Mus muscu
1.6 10989 10 AF45013884 AF450141 Mus muscu
1.6 41037 2 AC017476 Drosophil
1.6 50295 6 AX695884 Sequence
1.6 99416 9 AC105276 Homo sapi
1.6 110000 2 AC110828_4 Continuation (5 of
1.6 110000 2 AC110828_5 Continuation (6 of
1.6 114252 5 BX649441 Zebrafish

ALIGNMENTS

758 1373 bp DNA linear PAT 27-AUG-2002
or necrosis factor related ligand.
758
758.1 GI:22608361
11505407-A/2.
ified
ified.
ases 1 to 1373)
sportliche,X. and Browning J.L.
or necrosis factor related ligand
t: JP 2001505407-A 2 24-APR-2001;
N INC.THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GENEVA
INF family related protein
JP 2001505407-A/2
24-APR-2001
07-AUG-1997 JP 1998508239
07-AUG-1996 US 60/023541,18-OCT-1996 US 60/028515 PR
R-1997 US 60/040820
YVES CHICHEPORTICHE,JEFFREY L BROWNING
C12N15/28,C07K14/525,G01N33/68,C07K16/24,C12N15/11,A61K48/00,
C12N5/10,
A61K39/395,A61K38/19,C07K14/705,C12N15/12
Strandedness: Double;
Topology: Linear;
Key Location/Qualifiers
CDS 1..852.
Location/Qualifiers
1..1373
/organism="unidentified"
/mol_type="genomic DNA"

ORIGIN
/db_xref="taxon:32644"
Query Match 100.0%; Score 1373; DB 6; Length 1373;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1373; Conservative 0; Mismatches 0; Indels 0; C
QY 1 ATGTCAATTTAGTACTTTGAAATTTCCGCCCGCGGCTCCCTCCCGGATCCCT
Db 1 ATGTCAATTTAGTACTTTGAAATTTCCGCCCGCGGCTCCCTCCCGGATCCCT
QY 61 TCCCGGAGTGGGGGGGGTGTAGGAGGACACAGCCCCCGCCCCCATGGCGCCCC
Db 61 TCCCGGAGTGGGGGGGGTGTAGGAGGACACAGCCCCCGCCCCCATGGCGCCCC
QY 121 AGCCAGAGGCGGAGGCGCGCGGGGAGCGCGGCGACCGCGCTGCTGGTCCCGCT
Db 121 AGCCAGAGGCGGAGGCGCGCGGGGAGCGCGGCGACCGCGCTGCTGGTCCCGCT
QY 181 CTGGGCTTGGGCTGGCGCTGGCGCTGCTGGCGCTTCTGTGGCGGTGGTCAATT
Db 181 CTGGGCTTGGGCTGGCGCTGGCGCTGCTGGCGCTTCTGTGGCGGTGGTCAATT
QY 241 AGCCGGGATCGTGTCCGCCAGAGAGCTGCCCCAGGAGGAGCTGGTGGCAGAGG
Db 241 AGCCGGGATCGTGTCCGCCAGAGAGCTGCCCCAGGAGGAGCTGGTGGCAGAGG
QY 301 CAGGACCCGCTGGAACTGAAATCCCCCAGACAGAAAGCCAGATCTCTGGCGCTT
Db 301 CAGGACCCGCTGGAACTGAAATCCCCCAGACAGAAAGCCAGATCTCTGGCGCTT
QY 361 AACCCAGTAGTTGGCGCTCGCAGAGTGCACATAAGGCGGAAACACAGGCGTC
Db 361 AACCCAGTAGTTGGCGCTCGCAGAGTGCACATAAGGCGGAAACACAGGCGTC
QY 421 GCGATCGAGCCCATATGAGTTTCATCCAGACCTGGACAGAGCGGAGCGCAGG
Db 421 GCGATCGAGCCCATATGAGTTTCATCCAGACCTGGACAGAGCGGAGCGCAGG
QY 481 GTGGACGGGACAGTGTGGCTGGGAGGAAGCAGAAATCAACAGCTCCAGCCCTC
Db 481 GTGGACGGGACAGTGTGGCTGGGAGGAAGCAGAAATCAACAGCTCCAGCCCTC
QY 541 TACAACCGCCAGATCGGGAGTTTATAGTACCCGGGCTGGGCTTACTACTGT
Db 541 TACAACCGCCAGATCGGGAGTTTATAGTACCCGGGCTGGGCTTACTACTGT
QY 601 CAGGTGCATTTGATGAGGGGAGGCTGTCTACCTGAAGCTGACCTTGTGGTGG
Db 601 CAGGTGCATTTGATGAGGGGAGGCTGTCTACCTGAAGCTGACCTTGTGGTGG
QY 661 GTGCTGGCCCTCGCTGCTGGAGGAATTTCTCAGCCACTGCGGCGCAGTTCCCTCG
Db 661 GTGCTGGCCCTCGCTGCTGGAGGAATTTCTCAGCCACTGCGGCGCAGTTCCCTCG
QY 721 CAGTCCGCTTGCACAGTGTCTGGGCTGTGGGCTGTGGGCTGGGCGGAGGCTCCC
Db 721 CAGTCCGCTTGCACAGTGTCTGGGCTGTGGGCTGTGGGCTGGGCGGAGGCTCCC
QY 781 ATCCGCAACCTCCCTGGGCGCCATCTCAAGGCTGCCCCCTTCTCACTACTTCG
Db 781 ATCCGCAACCTCCCTGGGCGCCATCTCAAGGCTGCCCCCTTCTCACTACTTCG
QY 841 TTCCAGGTTCACTGAGGGGCGCTGTCTCCCAACAGTGTCCAGGCTGCGGCT
Db 841 TTCCAGGTTCACTGAGGGGCGCTGTCTCCCAACAGTGTCCAGGCTGCGGCT
QY 901 CGACAGCTCTGTGGGCAACCGGCTCCCTCTGCCCCACCTCAGCCGCTCTTTGCT
Db 901 CGACAGCTCTGTGGGCAACCGGCTCCCTCTGCCCCACCTCAGCCGCTCTTTGCT
QY 961 CTGTGCCCCCTCTCTAGAGGCTGCGCTGGGCTGTCTCAGGTGTTTTCATCCCA
Db 961 CTGTGCCCCCTCTCTAGAGGCTGCGCTGGGCTGTCTCAGGTGTTTTCATCCCA

06:25:14 2004

us-09-245-198a-3.oligo.rge

TCTCTGAGCTCCCTCCCTGGCCAGACACCCACGGGCAATTGTCTTCACTGTACTCTGTG 1080
 AAGGATGGGTCCAGAAGACCCCACTTTCAGGCACCTAAGAGGGGCTGGACCTGCGCGCA 1228
 AAGGATGGGTCCAGAAGACCCCACTTTCAGGCACCTAAGAGGGGCTGGACCTGCGCGCA 1140
 AGCCAAAGAGACTGGGCTCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1288
 AGCCAAAGAGACTGGGCTCTAGGCCAGGAGTTCCCAAATGTGAGGGGCGAGAAACAAG 1200
 AGCTCCTCCCTTGAGAAATTCCTCTGTGGATTTTAAACAGATATATTTTATTATT 1348
 AGCTCCTCCCTTGAGAAATTCCTGTGGAATTTTAAACAGATATATTTTATTATT 1260
 GTGACAAAATGTTGATAATGG 1373
 GTGACAAAATGTTGATAATGG 1285

324
324
324
324.1
324.1

ence 3 from Patent WO0153486.
GI:15391154

1353 bp DNA linear PAT 30-AUG-2001

sapiens (human)
sapiens
yocota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kazuki A. J., Goddard, A., Godowski, P. J., Gurney, A. L.,
Pan, J., Pitti, R. M., Roy, M. A., Smith, V.,
D. M., Matanabe, C. K. and Wood, W. I.
Associations and methods for the treatment of tumour
t: WO 0153486-A 3 26-JUL-2001;

```

Location/Qualifiers
1. .1353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

```

85.4%; Score 1172; DB 6; Length 1353;
larity 99.8%; Pred. No. 0;
Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

TCCCTCGGATCCCGGATGGGGGGGGCGGTGAGGAGGCACAGCCCCCGGCCCATG 108
 TCCCTCGGATCCCGGATGGGGGGGGCGGTGAGGAGGCACAGCCCCCGGCCCATG 60 -
 GCGCGTCGAGCCAGAGCGCGAGGGGGCGCCGGGGGAGCCGGGCACCGCCCTCGTG 168
 CCGCGTCGAGATCAGAGGGCGAGGGGGCGCGGGGGAGCCGGGCACCGCCCTCGTG 120

[illegible]

GAGGAGGAC CAGGACCCGTAATCCCGACAGAGAAAGCCGAGATCCT 348
 GAGGAGGAC CAGGACCCGTAATCCCGACAGAGAGAAAGCCGAGATCCT 300
 CCCCTTTCTCTGAACCGACTAGTTTCGGCCCTCGCAGAAAGTCACCTTAAGGGCCGAGAAACA 408
 CCCCTTTCTCTGAACCGACTAGTTTCGGCCCTCGCAGAAAGTCACCTTAAGGGCCGAGAAACA 360

QY	CGGGCTCGAAGAGCGATCGCAGCCCATATGATGAAGTTTCATCCACGACCTTGGACAGG	
Db	361	CGGGCTCGAAGAGCGATCGCAGCCCATATGATGAAGTTTCATCCACGACCTTGGACAGG
QY	469	CGCGAGGCAGGTGTGGACGGGACAGTGAAGTGGCTGGGAGGAAGCCAGAAATCAACA
Db	421	CGCGAGGCAGGTGTGGACGGGACAGTGAAGTGGCTGGGAGGAAGCCAGAAATCAACA
QY	529	AGCCCTCTGGCTACAAACGCCAGATCGGGGAGTTTATAGTCAACCCGGGCTGGGC
Db	481	AGCCCTCTGGCGCTAACAAACCAGATCGGGGAGTTTATAGTCAACCCGGGCTGGGC
QY	589	TACCTGTACTGTCAAGTGCACCTTGTATGAGGGAAGGCTGTCTACTGAAGCTGG
Db	541	TACCTGTACTGTCAAGTGCACCTTGTATGAGGGAAGGCTGTCTACTGAAGCTGG
QY	649	CTGGTGGATGCTGTGGCCCTGGCTGCCTCGAGGGAATTCCTCAGCCACTCGGG
Db	601	CTGGTGGATGCTGTGGCCCTGGCTGCCTCGAGGGAATTCCTCAGCCACTCGGG
QY	709	TCCCTCGGGCCCCAGCTCCGCTCTGCGAGGTGTCTGGGCTGTGTGGCCCTCGGGC
Db	661	TCCCTCGGGCCCCAGCTCCGCTCTGCGAGGTGTCTGGGCTGTGTGGCCCTCGGGC
QY	769	TCCTCCCTGGCGATCCGCAACCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCC
Db	721	TCCTCCCTGGCGATCCGCAACCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCC
QY	829	TACTTTCGGACTCTTCCAGGTTCACTGAGGGGGCCCTGGTCTCCCCACAGTGTCCC
Db	781	TACTTTCGGACTCTTCCAGGTTCACTGAGGGGGCCCTGGTCTCCCCCGAGTGTCCC
QY	889	GCGGCTCCGCTGCACAGCTCTCTGGGCAACCGGTCCGCTCTGCCCCACCCCTCAG
Db	841	GCGGCTCCGCTGCACAGCTCTCTGGGCAACCGGTCCGCTCTGCCCCACCCCTCAG
QY	949	CTTTGCTCCAGACCTTGCCCTTCCCTCTAGAGGCTGCTGGGCTGTGTACAGTGT
Db	901	CTTTGCTCCAGACCTTGCCCTTCCCTCTAGAGGCTGCTGGGCTGTGTACAGTGT
QY	1009	TGCCACATAAATACAGTATTCGCACTCTTATCTTACAACTCCGCCACGCCCACT
Db	961	TGCCACATAAATACAGTATTCGCACTCTTATCTTACAACTCCGCCACGCCCACT
QY	1069	CTCTCACTAGCTCCCCAAATCCCTGACCCCTTTGAGGCCCCCAAGTGAATCTCGACTCCC
Db	1021	CTCTCACTAGCTCCCCAAATCCCTGACCCCTTTGAGGCCCCCAAGTGAATCTCGACTCCC
QY	1129	GCACAGACCCCGAGGGATTTGTGTCTACTGTACTCTGTGGGCAAGGATGGGTCC
Db	1081	GCACAGACCCCGAGGTCAITTTGTGTTCACCTGTACTCTGTGGGCAAGGATGGGTCC
QY	1189	ACCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGAC
Db	1141	ACCCCACTTCAGGCACCTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGAC
QY	1249	CTAGGCCAGGATTCGCCAAATGTGAGGGGGCGAGAAACAAGCAAGCTCTCCCTCT
Db	1201	CTAGGCCAGGATTCGCCAAATGTGAGGGGGCGAGAAACAAGCAAGCTCTCCCTCT
QY	1309	TTCCCTGTGGATTTTTTAAAAACAGATATTTATTTTATTTATTTGTGACAAAAATGT
Db	1261	TTCCCTGTGGATTTTTTAAAAACAGATATTTTTTTTTATTTATTTGTGACAAAAATGT
QY	1369	AATGG 13173
Db	1321	AATGG 1325

RESULT 4	LOCUS	AY358870	1353 bp	mRNA	linear	PRI
AY358870	DEFINITION	Homo sapiens clone DNA30879 Apo-3 L3-gand (UNQ181)	mRNA, c			

```
10 0.1 GI:37182857
11 1A.
12 sapiens (human)
13 Meta: Chordata; Craniata; Vertebrata; Euteleostomi;
14 ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
15 ses 1 to 1353
16 i.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
17 ., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,
18 ., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
19 ., S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
20 ., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
21 ., J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
22 ., J., Vagts,A., Vanden,K., Watanabe,C., Wleand,D., Woods,K.,
23 ., A., Wood,W.I. and Godowski,P.
24 uted Protein Discovery Initiative (SPDI), a Large-Scale
25 to Identify Novel Human Secreted and Transmembrane Proteins:
26 nformatics Assessment
27 Res. 13 (10), 2265-2270 (2003)
28 09
29 ses 1 to 1353
30 i.F.
31 Submission
32 ted (01-AUG-2003) Department of Bioinformatics, Genentech,
33 1 DNA Way, South San Francisco, CA 94080, USA
34 Location/Qualifiers
35 1. .1353
36 /organism="Homo sapiens"
37 /mol_type="mRNA"
38 /db_xref="taxon:9606"
39 /clone="DNA30879"
40 1. .1353
41 /locus_tag="UNQ181"
42 /locus_tag="UNQ181"
43 /note="PRO207"
44 /codon_start=1
45 /product="Apo-3 Ligand"
46 /protein_id="AAQ89229.1"
47 /db_xref="GI:37182858"
48 /translation="MAARRSRRRRGPGTALLPLALGLALACILGLAVSL
49 GSRASLAEPAQELVAEPDSELPNTPQESODPAPFLNLRVPRSPKGRKTR
50 ARRALAHYVHPRPQDGAQAGVDGTVSGHEEARINSSPLRVNRIQGEIVTEAGL
51 YLYCVHFDGKAVYLLKDLVLLDGLVLRCLREFSATAASLSLGPQLRCLQVSGLLAL
52 RFGSSLRITRTLPWAHLKAAPFLTVFGLFQVH"
53
54 85.4%; Score 1172; DB 9; Length 1353;
55 arity 99.8%; Pred. No. 0;
56 onservative 0; Mismatches 3; Indels 0; Gaps 0;
57
58 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 108
59 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 60
60 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 168
61 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 120
62 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 228
63 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 180
64 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 288
65 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 240
66 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 348
67 CCCTCGGTCCTCCGGATGGGGGGCGGTGAGGCAGGCACACCCCGCCCGCCCATG 300
```

06:25:14 2004

us-09-245-198a-3.oligo.rge

872 sapiens Apo3/DR3 ligand (APO3L) mRNA, complete cds.
872 sapiens (human)
872.1 GI:3108230
355 Biol. 8 (9), 525-528 (1998)
43 ases 1 to 1368)
ases,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and
nazi,A.
t Submission
ted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,
San Francisco, CA 94080, USA
Location/Qualifiers
1..1368
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
1..1368
/gene="APO3L"
58..807
/gene="APO3L"
/function="binds to the death-domain-containing receptor
Apo3/DR3"
/note="TNF homolog"
/codon_start=1
/product="Apo3/DR3 ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAAPRRRRRGEPTALLPLALGLHALACIGILLAVVSL
GSRASLNAOEPALAEVDEPDSQDPAPFLNLRVRRSPAPKGRKTR
ARRALAHYVHPRGQAGAGVDGTVSGWEARINSSEPLRNQIGHFIVTRAGL
YLYLCQVHFDGKAVYLLKLDLVDGLALRCLEEFATAAASSLGPQLRLCQVSLAL
RPGSSLRITLPLWAHLKAAPELTYFLGFQVH"
85.4%; Score 1172; DB 9; Length 1368;
Larity 99.8%; Pred.No.0;
Conservative 0; Mismatches 3; Indels 0; Gaps 0;
ATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGAGGACACAGCCCCCCCCCATG 108
ATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGAGGACACAGCCCCCCCCCATG 60
GCCCTCGGAGCCAGAGCGAGAGGGGGCGCGGGGGAGAGCGGGGACCGCCCTGCTG 168
GCCCTCGGAGCCAGAGCGAGAGGGGGCGCGGGGGAGAGCGGGGACCGCCCTGCTG 120
GCCCTCGGAGCCAGAGCGAGAGGGGGCGCGGGGGAGAGCGGGGACCGCCCTGCTG 228
GCCCTCGGAGCCAGAGCGAGAGGGGGCGCGGGGGAGAGCGGGGACCGCCCTGCTG 180
GCCCTCGGAGCCAGAGCGAGAGGGGGCGCGGGGGAGAGCGGGGACCGCCCTGCTG 288
GCCCTCGGAGCCAGAGCGAGAGGGGGCGCGGGGGAGAGCGGGGACCGCCCTGCTG 240
AGAGGAGGACAGAGCCCGTGGAACTGAAATCCCCAGACAGAAAGCCAGGATCCT 348

241 GCAGAGGAGGACCCAGGACCCGTCGGAACTGAATCCCCAGACAGAAAGAGCCAGG
349 GGGCCCTTTCCTGAACCGAGCTAGTTGGCTCGCAGAAAGTGCACTAAAGCCCGGA
301 GGGCCCTTTCCTGAACCGAGCTAGTTGGCTCGCAGAAAGTGCACTAAAGCCCGGA
409 CGGGCTCGAAGAGCGATCGACGCCCAATTATGAAGTTTCATCCAGACTGGACAGG
361 CGGGCTCGAAGAGCGATCGACGCCCAATTATGAAGTTTCATCCAGACTGGACAGG
469 GGGCAGGAGGTTGTCAGCGGACAGTGGCTGGGAGGAGCCAGAAATCAACA
421 GGGCAGGAGGTTGTCAGCGGACAGTGGCTGGGAGGAGCCAGAAATCAACA
529 AGCCCTCTCGGCTACAAACCGCAGATCGGGAGTTTATAGTCAACCGGGCTGGGC
481 AGCCCTCTCGGCTACAAACCGCAGATCGGGAGTTTATAGTCAACCGGGCTGGGC
589 TACCTGTACTGTGAGTGCACTTTTGTATGAGGGGAGAGGCTGTCTACCTGAAGCTGG
541 TACCTGTACTGTGAGTGCACTTTTGTATGAGGGGAGAGGCTGTCTACCTGAAGCTGG
649 CTGGTGGAGTGTGCTGGCCCTCGCTGCTGAGGAGAAATTTCTCAGCCACTGGCG
601 CTGGTGGAGTGTGCTGGCCCTCGCTGCTGAGGAGAAATTTCTCAGCCACTGGCG
709 TCCTCTGGGGCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGCCCTTCGCGC
661 TCCTCTGGGGCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGCCCTTCGCGC
769 TCCTCTGGGGCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGCCCTTCGCGC
721 TCCTCTGGGGCCCCAGCTCCGCTCTGCGAGTGTCTGGGCTGTGGCCCTTCGCGC
829 TACTTCGGAGTCTTCCAGGTTTCACTGAGGGGGCCCTGCTCTCCACAGTGTGTC
781 TACTTCGGAGTCTTCCAGGTTTCACTGAGGGGGCCCTGCTCTCCACAGTGTGTC
889 GCGGGCTCCCTCCGACAGCTCTCTGGGACCCCGTCTCCCTCTGCCCCACCTCAAC
841 GCGGGCTCCCTCCGACAGCTCTCTGGGACCCCGTCTCCCTCTGCCCCACCTCAAC
949 CTTTGTCTCCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTGTCAGTGT
901 CTTTGTCTCCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTGTCAGTGT
1009 TCCACATATATACAGTATCCCACTCTTATCTTACAACTCCCCCAGCGCCAC
961 TCCACATATATACAGTATCCCACTCTTATCTTACAACTCCCCCAGCGCCAC
1069 CCTCCTAGCTTCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGTCTGACTCC
1021 CCTCCTAGCTTCCCAATCCCTGACCCCTTTGAGGCCCCCAGTGTCTGACTCC
1129 GCGACAGACCCCGAGGAGTGTGTTCACTGTACTCTGTGGGAGAGTGGGTCT
1081 GCGACAGACCCCGAGGCTCATGTGTCTCACTGTACTCTGTGGGAGAGTGGGTCT
1189 ACCCACTTCCAGGCACTAAGAGGGGCTGGACCTGGCGGAGAGGAGCCAAAGAGAG
1141 ACCCACTTCCAGGCACTAAGAGGGGCTGGACCTGGCGGAGAGGAGCCAAAGAGAG
1249 CTAGGCGAGGAGTTCCCAATGTGAGGGGGAGAGAAACAAGCAAGCTCTCCCTCT
1201 CTAGGCGAGGAGTTCCCAATGTGAGGGGGAGAGAAACAAGCAAGCTCTCCCTCT
1309 TTCCCTGTGGATTTTAAACAGATATATTTTATTTATTTATTTATTTGACAAATG
1261 TTCCCTGTGGATTTTAAACAGATATATTTTATTTATTTATTTATTTGACAAATG
1369 AATGG 1373
1321 AATGG 1325

EST	2	1421 bp	DNA	linear	PAT 27-AUG-2002
EF01	igand polypeptide.				
EF02	2				
EF03	2.1 GI:22636562				
EF04	522584-A/1.				
EF05	piens (human)				
EF06	piens				
EF07	ata; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
EF08	a; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
EF09	es 1 to 1421)				
EF10	izi.A.J., Marsters,S.A. and Pitti,R.				
EF11	igand polypeptide				
EF12	JP 2001522584-A 1 20-NOV-2001;				
EF13	ICH INC				
EF14	mno sapiens (human)				
EF15	, 2001522584-A/1				
EF16	, 1-NOV-2001				
EF17	, 1-OCT-1998 JP 2000516042				
EF18	, 1-OCT-1997 US 60/062037, 17-DEC-1997 US 60/069862 PI				
EF19	ASHKENAZI, SCOT A MARSTERS, ROBERT PITTI				
EF20	.2N15/09,A61K39/00,C07K14/705,C07K16/24,C12N15/00,A61K37/02				
EF21	o-3 ligand polypeptide				
EF22	y Location/Qualifiers				
EF23	1. .1421				
EF24	Location/Qualifiers				
EF25	1. .1421				
EF26	/organism="Homo sapiens"				
EF27	/mol_type="genomic DNA"				
EF28	/db_xref="taxon:9606"				
EF29	85.4%; Score 1172; DB 6; Length 1421;				
EF30	arity 99.8%; Pred. No. 0;				
EF31	nservative 0; Mismatches 3; Indels 0; Gaps 0;				
EF32	TCCTCGGTCCCGGATGGGGGGCGGTGAGGCGGACAGCCCCCGCCCCCATG 108				
EF33	TCCTCGGTCCCGGATGGGGGGCGGTGAGGCGGACAGCCCCCGCCCCCATG 94				
EF34	TCCTCGGAGCCAGAGCGGGGCGCGGGGGGAGCGGGGACACCGCCCTGCTG 168				
EF35	TCCTCGGAGCCAGAGCGGGGCGCGGGGGGAGCGGGGACACCGCCCTGCTG 154				
EF36	TCCTCGGTGGGCGTGGCCCTGGCCCTGGCCCTGGCCCTGGCCCTGGCCCTG 228				
EF37	TCCTCGGTGGGCGTGGCCCTGGCCCTGGCCCTGGCCCTGGCCCTGGCCCTG 214				
EF38	TTTGGGAGCCGGGCGATCGCTGTCCGCCCCAGGAGCTGCTCCAGGAGGAGCTGGT 288				
EF39	TTTGGGAGCCGGGCGATCGCTGTCCGCCCCAGGAGCTGCTCCAGGAGGAGCTGGT 274				
EF40	AGGAGACACAGACCCGTGGAACTGAATCCCGACAGAGAAAGCCAGGATCCT 348				
EF41	AGGAGACACAGGACCCGTGGAACTGAATCCCGACAGAGAAAGCCAGGATCCT 334				
EF42	CTTTCCTGAACCGACTAGTTCGGGCTCGCAGAAAGTGACCTAAAGCGCGGAAACA 408				
EF43	CTTTCCTGAACCGACTAGTTCGGGCTCGCAGAAAGTGACCTAAAGCGCGGAAACA 394				
EF44	CTCGAAGAGCGATCGCAGCCCATATGAAATTATCCAGCTTGGACAGGACGGA 468				
EF45	CTCGAAGAGCGATCGCAGCCCATATGAAATTATCCAGCTTGGACAGGACGGA 454				
EF46	AGGACAGGTGGACGGACAGTGAAGTGGTGGGAGGAGCCAGATCAACAGCTCC 528				
EF47	AGGACAGGTGGACGGACAGTGAAGTGGTGGGAGGAGCCAGATCAACAGCTCC 514				
EF48	CTCTCGGCTACAAACCGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTCTAC 588				

```

e
: US 6207642-A 1 27-MAR-2001;
  Location/Qualifiers
    1..1236
    /organism="unknown"
    /mol_type="unassigned DNA"

      69.8%; Score 958; DB 6; Length 1236;
      arity 99.6%; Pred. No. 0;
      onservative 0; Mismatches 5; Indels 0; Gaps 0;

      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 188
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 83
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 248
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 143
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 308
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 203
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 368
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 263
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 428
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 323
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 488
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 383
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 548
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 443
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 608
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 503
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 668
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 563
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 728
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 623
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 788
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 683
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 848
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 743
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 908
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 803
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 968
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 863
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 1028
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 923
      IAGGGGCGCCGGGGAGCCGGGACCCCGGACCCCGCTGCTGGTCCCGCTGGGCT 1088

```

```

Db 924 CCCACTTTATCTTACAAACCCCAACCGCCACTCTCCACTCACTAGCTCCCG
QY 1089 CTGACCCCTTTGAGGCCCCCAGTGTCTGACTCCCGCTGGCCACAGACCCCGAG
Db 984 CTGACCCCTTTGAGGCCCCCAGTGTCTGACTCCCGCTGGCCACAGACCCCGAG
QY 1149 TGTGTTCTACTTACTTGTGGGCAAGATGGTCCAGAGACCCCACTTCAGGCA
Db 1044 TGTGTTCTACTTACTTGTGGGCAAGATGGTCCAGAGACCCCACTTCAGGCA
QY 1209 AGGCGCTGAGCTGGCGCAGGAAAGCAAGAGACTGGGCTTAGGCCAGGAGTTCC
Db 1104 AGGCGCTGAGCTGGCGCAGGAAAGCAAGAGACTGGGCTTAGGCCAGGAGTTCC
QY 1269 TGTGAGGGGGCGAGAAACAAAGACAAAGCTCCTCCCTTGAGAAATTCCTGTGATTTT
Db 1164 TGTGAGGGGGCGAGAAACAAAGACAAAGCTCCTCCCTTGAGAAATTCCTGTGATTTT
QY 1329 CAGATATTATTTT 1341
Db 1224 CAGATATTATTTT 1236

RESULT 8
BD057124 1236 bp DNA linear PAT 2
LOCUS Member of the tnfr family useful for treatment and diagnos
DEFINITION disease.
ACCESSION BD057124
VERSION BD057124.1 GI:22602730
KEYWORDS JP 2001513626-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1236)
AUTHORS Wiley, S.R.
TITLE Member of the tnfr family useful for treatment and diagnos
JOURNAL Patent: JP 2001513626-A 1 04-SEP-2001;
COMMENT PN JP 2001513626-A/1
PD 04-SEP-2001
PF 12-FEB-1998 JP 1998535077
PR 12-FEB-1997 US 08/798692, 10-FEB-1998 US 09/021
PC C12Q1/68, C12N15/63, C07K14/525, C07K16/28, G01N33/50, G0
PC A61K38/19
CC Strandedness: Single;
CC Topology: Linear;
CC Key Location/Qualifiers.
FEATURES
source
  1..1236
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"

ORIGIN
Query Match 69.8%; Score 958; DB 6; Length 1236;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 5; Indels 0;

QY 129 GCGAGAGGGGGCGCGGGGGAGCGCGGACCGCCCTGCTGGTCCCGCTCGCGCTG
Db 24 GCGAGAGGGGGCGCGGGGGAGCGCGGACCGCCCTGCTGGTCCCGCTCGCGCTG
QY 189 GGGCGCTGGCGCTGGCGCTGCTGGCGCTCCTGCTGGCGCTGGTTCAGTTTGGGGAGC
Db 84 GGGCGCTGGCGCTGGCGCTGCTGGCGCTCCTGCTGGCGCTGGTTCAGTTTGGGGAGC
QY 249 ATCGCTGTCCCGCCAGGAGCTTCGCCAGGAGGAGCTGGTGGCAGAGGAGGACAG
Db 144 ATCGCTGTCCCGCCAGGAGCTTCGCCAGGAGGAGCTGGTGGCAGAGGAGGACAG

```

1 AACTGAATCCCGACAGAGAAAGCCAGGATCCCTGGCGCTTTCTCTGAACGACT 368
 2 AACTGAATCCCGACAGAGAAAGCCAGGATCCCTGGCGCTTTCTCTGAACGACT 263
 3 GCGCTCGCAGAGATGACCTAAAGGCGGAAACACAGGCTCGAAGAGCGATCGC 428
 4 GCGCTCGAAGAGTGCACCTAAAGGCGGAAACACAGGCTCGAAGAGCGATCGC 323
 5 CATATGAAGTTTATCAGCACTGACAGAGCGGAGCGGAGCGGAGCGGATGCGACGG 488
 6 CATATGAAGTTTATCAGCACTGACAGAGCGGAGCGGAGCGGAGCGGATGCGACGG 383
 7 CTGAGTGCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548
 8 CTGAGTGCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 443
 9 ATCGGAGGATTTATGATACCCCGGCTGGGCTCTACTACTGATCTGATGATGATGCA 608
 10 ATCGGAGGATTTATGATACCCCGGCTGGGCTCTACTACTGATCTGATGATGATGCA 503
 11 GATGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
 12 GATGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 563
 13 CGCTGCTGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 728
 14 CGCTGCTGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 623
 15 TGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 788
 16 TGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 683
 17 CCCTGGGCGCATCTCAAGGCTGCGGCTTCTCTCACTACTTCTGATCTTCTGAGGT 848
 18 CCCTGGGCGCATCTCAAGGCTGCGGCTTCTCTCACTACTTCTGATCTTCTGAGGT 743
 19 TGAGGGCGCTGCTGCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908
 20 TGAGGGCGCTGCTGCTGCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803
 21 GGGCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 968
 22 GGGCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 863
 23 TCTAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1028
 24 TCTAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
 25 CTCTTATCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1088
 26 CTCTTATCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
 27 CCCTTTGAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1148
 28 CCCTTTGAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1043
 29 TTCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208
 30 TTCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
 31 TCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1268
 32 TCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
 33 TCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1328
 34 TCTACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1223
 35 TATTATTTT 1341
 36 TATTATTTT 1236

RESULT 9
 BC019047
 LOCUS
 DEFINITION
 Homo sapiens tumor necrosis factor (ligand) superfamily, m
 transcript variant 2, mRNA (cdna clone MGC:20669 IMAGE:476
 complete cds.
 BC019047
 BC019047.2 GI:33874709
 MGC.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1642)
 Strausberg, R.L., Collins, F.S., Wagner, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schu
 Altschul, R.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bha
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heie
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Urdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Wallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Mye
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.F
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-1
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1642)
 Strausberg, R.
 Direct Submission
 Submitted (07-DEC-2001) National Institutes of Health, Man
 Gene Collection (MGC), Cancer Genomics Office, National Cc
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:17512138
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramur
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhgri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Broc
 Districh, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legai
 Maduro, Q.I., Masiello, C., Maskeri, B., Mastrian, S.D., McClot
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touci
 Tsurgenc, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wigglir
 Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.l>
 Series: IRAL Plate: 30 Row: p Column: 5
 This clone was selected for full length sequencing because
 passed the following selection criteria: matched mRNA gi:
 Location/Qualifiers
 1. 1642

FEATURES
 source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20669 IMAGE:4766071"
/tissue_type="Primary B-Cells from Tonsils"
/clone_lib="NIH MGC 48"
/lab_host="DH10B-R-"
/note="vector: pOTB7"
1. .1642
/gene="TNFSF12"
/note="synonyms: TWEAK, DR3LG, APO3L, MGC20669"
/db_xref="LOCusID: 8742"
/db_xref="MIM:602695"
97..501
/codon_start=1
/product="cumor necrosis factor (ligand) superfamily,
member 12, isoform 2"
/protein_id="AAH19047.1"
/db_xref="GI:17512139"
/db_xref="JcucID:18742"
/translaton="MAARRRRRRRRRGEPTVLLVPIALGILGALACIGLLAVVLS
GSRASLGAQEPAQELVAEDQDSLELNPQTEESQDPAFFNRLVRRPSAPKGRKTR
ARRAIAAHVEVHPRPQDGAQDGGYTTCLR"
58.0%; Score 796; DB 9; Length 1642;
arity 99.8%; Pred. No. 0;
conservative 0; Mismatches 2; Indels 0; Gaps 0;
.TGTGGACGGACAGTGAGTGCTGGAGGAAGCCAGAAATCAACAGCTCCAGCCCTC 535
.TGTGGACGGACAGTGAGTGCTGGAGGAAGCCAGAAATCAACAGCTCCAGCCCTC 765
.CTCAACCCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGT 595
.CTCAACCCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACCTGT 825
.TCAGGTGCACTTTGATGAGGGGAAGCGTGCTACCTGAAGCTGGAATTGCTGGTGG 655
.TCAGGTGCACTTTGATGAGGGGAAGCGTGCTACCTGAAGCTGGAATTGCTGGTGG 885
.TGTGCTGGCCCTCGCTGCGAGGAATTTCTACGCCACTGCGGCCAGGTTCCTCG 715
.TGTGCTGGCCCTCGCTGCGAGGAATTTCTACGCCACTGCGGCCAGGTTCCTCG 945
.TCAGCTCGCCCTCTGCCAGGTGTCTGGGCTGTGTGGCCCTCGGGCCAGGCTCTCC 775
.TCAGCTCGCCCTCTGCCAGGTGTCTGGGCTGTGTGGCCCTCGGGCCAGGCTCTCC 1005
.TGATCGGCACCTTCGCTGGGCCCATCTCAAGGCTGCCCTTCCTCACCCTACTTCG 835
.TGATCGGCACCTTCGCTGGGCCCATCTCAAGGCTGCCCTTCCTCACCCTACTTCG 1065
.TCTCCAGGTTCACTGAGGGGCCCTGGTCTCCGCCACAGTCGTCCAGGCTGCCGGCT 895
.TCTCCAGGTTCACTGAGGGGCCCTGGTCTCCGCCAGTCGTCCAGGCTGCCGGCT 1125
.TCGACAGCTCTCTGGGGCACCCGGTCCCTCTGTGCCCCACCCCTCAGCGGCTCTTTGCT 955
.TCGACAGCTCTCTGGGGCACCCGGTCCCTCTGTGCCCCACCCCTCAGCGGCTCTTTGCT 1185
.TACCTGCCCTTCCTCTAGAGGCTGCTGGGCCGTGTTCAGTGTGTTTCCATCCACA 1015
.TACCTGCCCTTCCTCTAGAGGCTGCTGGGCCGTGTTCAGTGTGTTTCCATCCACA 1245
.ATACAGTATTGCCACTCTTATCTTTACAACTCCGCCACCGGCCCACTCTCCACCTCACT 1075
.ATACAGTATTGCCACTCTTATCTTTACAACTCCGCCACCGGCCCACTCTCCACCTCACT 1305
.TCCCCAATCCCTGACCCCTTTGAGGCCGCCAGTGAATCTGCACTCCCCCTGGCCACAG 1135
.TCCCCAATCCCTGACCCCTTTGAGGCCGCCAGTGAATCTGCACTCCCCCTGGCCACAG 1365

```

Qy	1136	ACCCCCAGGCGCATTGTTTCACTGTACTCTCTGGGCAAGGATGGTCCACGAAGACC	
Db	1366	ACCCCCAGGCGCATGTGTTCATCTGTACTCTGTGGGCAAGGATGGTTCAGAGACC	
Qy	1196	TTCAGGCACCTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGGCTTA	
Db	1426	TTCAGGCACCTAAGAGGGGCTGGACCTGGCGGCAGGAAGCCAAAGAGACTGGGGCTTA	
Qy	1256	AGGAGTTCCCAATGTGAGGGGCGAAGAACAGACAAGACTCTCTCCCTTGAGATTTC	
Db	1486	AGGAGTTCCCAATGTGAGGGGCGAAGAACAGACAAGACTCTCTCCCTTGAGATTTC	
Qy	1316	TGGATTTTAAAAACAGATATTATTTTATTATTATTGTGACAAAATGTTGATAAAT	
Db	1546	TGGATTTTAAAAACAGATATTATTTTATTATTATTGTGACAAAATGTTGATAAAT	
RESULT 10			
AXI180714			
LOCUS	AXI180714	898 bp	DNA linear PAT 06
DEFINITION	Sequence 1 from Patent WO0145730.		
ACCESSION	AXI180714		
VERSION	AXI180714.1	GI:15132570	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1 .		
AUTHORS	Wiley, S.R.		
TITLE	Weak receptor		
JOURNAL	Patent: WO 0145730-A 1 28-JUN-2001;		
FEATURES	IMMUNEX CORPORATION (US)		
source	Location/Qualifiers		
	1 . 898		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
CDS	52 . 873		
	/note="unnamed protein product; human TWEAK fusio		
	construct"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAC50004.1"		
	/db_xref="GI:15132571"		
	/db_xref="REMTREMBL:CAC50004"		
	/translation="MAGRTSLSLAFGLLCLPWLQESATSSDRMKQ:		
	KIVHENEIARIKKLIGERTSLRSLASLQAEPQELVAEEDDP:		
	QDPAPFLVRUPRAGKGRTRARRIAAHVEVHPRPQDGAQGV		
	RINSSPLRVNROIREFITRAGLYLVCOVHDFDKAVYLKLLVDX		
	FSATAASSLPQRLQCVGSLALRPSSLRIRTLPAWHLKAAPLFI		
ORIGIN			
	Query Match	45.5%;	Score 625; DB 6; Length 898;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 625; Conservative	0;	Mismatches 0; Indels 0;
Qy	232	AGTTTGGGAGCCGGGCATCGCTGTCCGCCAGGAGCTGCCCCAGGAGGAGCTGZ	
Db	250	AGTTTGGGAGCCGGGCATCGCTGTCCGCCAGGAGCTGCCCCAGGAGGAGCTGZ	
Qy	292	GAGGAGGACGAGACCCGCTCGGAACCTGAATCCCCACAGAGAAGACGAGATC	
Db	310	GAGGAGGACGAGACCCGCTCGGAACCTGAATCCCCACAGAGAAGACGAGATC	
Qy	352	CCTTCTTGAAACCGACTAGTTTCGGCTTCGAGAGTGCACCTAAAGGCGGAAAAA	
Db	370	CCTTCTTGAAACCGACTAGTTTCGGCTTCGAGAGTGCACCTAAAGGCGGAAAAA	
Qy	412	GCTCGAAGCGATCGCAGCCCATTTAGTTTCATCCACGACCTGGACAGGACG	
Db	430	GCTCGAAGCGATCGCAGCCCATTTAGTTTCATCCACGACCTGGACAGGACG	
Qy	472	CAGGCAGGTGTGGACGGGACAGTGTAGTGGCTGGGAGGAAGCCAAATCAAACAGT	

|||||
 AGGTGTGGA CGGACAGTGA GTGGCTGGGAGGAGCCAGATCA CACAGCTCCAGC 549
 |||||
 TCGCTCAACACCGCCAGATCGGGAGGTTTATAGTACACCGGGCTGGGCTTACTAC 591
 |||||
 TCGCTCAACACCGCCAGATCGGGAGGTTTATAGTACACCGGGCTGGGCTTACTAC 609
 |||||
 CTGTGAGTGCACCTTTTCATGAGGAGGCTGTCTACCTGAAGCTGGAGCTTGCTG 651
 |||||
 CTGTGAGTGCACCTTTTCATGAGGAGGCTGTCTACCTGAAGCTGGAGCTTGCTG 669
 |||||
 TGTGTGCTGCGCCCTGCGCTGCGCTGGAGGAATTCAGCCACTGCGCCAGTTC 711
 |||||
 TGTGTGCTGCGCCCTGCGCTGCGCTGGAGGAATTCAGCCACTGCGCCAGTTC 729
 |||||
 TCGCCAGCTGCGCTGCGCCAGTGTCTGGGCTGTGGCCCTGGGCGCAGGGTCC 771
 |||||
 TCGCCAGCTGCGCTGCGCCAGTGTCTGGGCTGTGGCCCTGGGCGCAGGGTCC 789
 |||||
 TCGCCAGCTGCGCCCTGCGCCAGTGTCTGGGCTGTGGCCCTGGGCGCAGGGTCC 831
 |||||
 TCGCCAGCTGCGCCCTGCGCCAGTGTCTGGGCTGTGGCCCTGGGCGCAGGGTCC 849
 |||||
 ACTCTTCAGGTTCACTGAG 856
 |||||
 ACTCTTCAGGTTCACTGAG 874

6 60268 bp DNA linear PRI 31-OCT-2002
 piens chromosome 17, clone RP11-186B7, complete sequence.

6.10 GI:24431829

piens (human)

piens
 ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Primates; Catarrhini; Hominidae; Homo.

es 1 to 60268

B., Nusbaum, C. and Lander, E.

piens chromosome 17, clone RP11-186B7

shed

es 1 to 60268

B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,
 Castle, A., Colling, R., Collins, S., Collymore, A.,
 DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
 a, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
 J., Gardyna, S., Grant, G., Hags, B., Heaford, A., Horton, L.,
 J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Y., J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
 P., McGurk, A., McKernan, K., McLaughlin, J., Melidrim, J.,
 J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 n, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 i., Ye, W.J., Zimmer, A. and Zody, M.

Submission
 ed (08-DEC-1999) Whitehead Institute/MIT Center for Genome
 h, 320 Charles Street, Cambridge, MA 02141, USA

es 1 to 60268

B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 i., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
 a, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 i, S., Gord, S., Graham, L., Grand-Pierre, N., Hags, B.,
 L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 i, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 MacLean, C., Macdonald, P., Major, J., Matthews, C.,
 y, M., Melidrim, J., Meneus, L., Mihova, T., Mlenga, V.,
 T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

TITLE JOURNAL

REFERENCE AUTHORS

O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rog
 Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Seve
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Tal
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.
 Viel, R., Vo, A., Wilson, J., Wu, X., Wyman, D., Young, G., Zainc
 Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (08-OCT-2002) Whitehead Institute/MIT Center for
 Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 60268)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Andersc
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter
 Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.
 Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge
 Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hags, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,
 Matthews, C., McCarthy, M., Melidrim, J., Meneus, L., Mihova, T.,
 Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann,
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham,
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zc

Direct Submission

Submitted (31-OCT-2002) Whitehead Institute/MIT Center for
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Oct 31, 2002 this sequence version replaced gi:23592141.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Rese

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3849

Center clone name: 186_B_7

Only the first 60.3 kilobases of this clone are being submi:
 The remainder overlaps accession number AC113189 [WICGR pri
 L23113].

FEATURES

source

Location/Qualifiers
 1. 60268
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"

/clone="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"

/clone_lib="RP11-186B7"


```
Query Match      45.2%; Score 620; DB 9; Length 60268;
Best Local Similarity 99.6%; Pred.No. 0;
Matches 770; Conservative 0; Mismatches 3; Indels 0; C
```

RESULT 12	AY081051	AY081051	1816 bp	mrna	linear	PRI 1
LOCUS						
DEFINITION						
COMMENT						
FEATURES						
ORIGIN						
BASES						
GC						
ORIGIN						
BASES						
GC						

1
1.1 GI:24934973
piens (human)
piens
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 1816)
Balade,B., Medema,J.P., Lopez-Praga,M., Lozano,J.C.,
oten,G.M., Picard,A., Martinez-A.C., Garcia-Sanz,J.A. and
genous hybrid mRNA encodes TWE-PRIL, a functional cell
TWEAK-APRIL fusion protein
21 (21), 5711-5720 (2002)
9
es 1 to 1816)
Balade,B., Garcia-Sanz,J.A. and Hahne,M.
Submission
ed (21-FEB-2002) DIO, CNB, ctra de Colmenar Viejo, MADRID,
28049, Spain
Location/Qualifiers
1. .1816
/organism="Homo sapiens"
/mol type="mRNA"
/db xref="taxon:9606"
/chromosome="17"
/notes="expressed in primary T lymphocytes and monocytes"
56. .1048
/note="fusion protein; TWEAK/APRIL hybrid; TWEAK and APRIL
are members of the TNF ligand family, corresponding to
TNFSF12 and TNFSF13, respectively"
/codon_start=1
/product="TWE-PRIL"
/protein_id="AA090443.1"
/db xref="GI:24934974"
/translation="MAARRSQRRRRGEPTALLVPLALGLALACLGLLAVSL
GSRALSQAQPAQELVAEDQPSLNPOTESQDPAPFLNLRPRRSPKGRKTR
ARRALAHVEHPREGDGAQGVGTSGWEARINSPLRNROIGSFIVTRAGL
YLLYQSSDALEAWENERSKRAVLTKQKHGSVLHLPINATSKDDSDVTENVW
OPALLRGGLQAGYGVRIQAGVYLLYSQVLFQDVTFTMGQVVSREGQRETLFRC
IRSPSPHPRAYNSCYAGVPHLQDILLSVILPRARAKNLSPHGLFLGFVKL"
40.3%; Score 553; DB 9; Length 1816;
arity 100.0%; Pred.No. 8.3e-297; Indels 0; Gaps 0;
conservative 0; Mismatches 0; Indels 0; Gaps 0;
CTCGGCTCCGGGATGGGGGGCGGTGAGGAGGACACAGCCCCCGCCCATGGC 110
CTCGGCTCCGGGATGGGGGGCGGTGAGGAGGACACAGCCCCCGCCCATGGC 60
CGTCGAGCCAGAGCGGAGGGGGCGGGGGGAGCGGCGACCGCCCTGCTGGT 170
CGTCGAGCCAGAGCGGAGGGGGCGGGGGGAGCGGCGACCGCCCTGCTGGT 120
CTCGGCTCCGGGATGGGGGGCGGTGAGGAGGACACAGCCCCCGCCCATGGC 230
CTCGGCTCCGGGATGGGGGGCGGTGAGGAGGACACAGCCCCCGCCCATGGC 180
TTGGGAGCCGCGGATCCCTGTCCGCCAGGAGCGCTCCCGAGGAGGAGCTGGTGGC 290
TTGGGAGCCGCGGATCCCTGTCCGCCAGGAGCGCTCCCGAGGAGGAGCTGGTGGC 240
XGAGGACCGAGGACCGCTCGGAACTGAATCCCGAGAGAGAAAGCAGGATCCTGC 350
XGAGGACCGAGGACCGCTCGGAACTGAATCCCGAGAGAGAAAGCAGGATCCTGC 300
TTTCCTGAACCGACTAGTTCGGCTCGCAGAGTGCACCTTAAGGCGGAAACACG 410
TTTCCTGAACCGACTAGTTCGGCTCGCAGAGTGCACCTTAAGGCGGAAACACG 360
TCGAAGAGCGATCGCAGCGCCATTATGAAGTTTCATCCAGCCTTGGACAGGAGCG 470

Db 361 GGCTCGAAGAGCGGATCGCAGCCCATTTAAGATTTCATCCAGACCTGGACAGCGCG
QY 471 GCAGGACAGGTGTGGACGGGACAGTGTAGTGGTGGAGGAGGAGCAATCAACAGCTC
Db 421 GCAGGACAGGTGTGGACGGGACAGTGTAGTGGTGGAGGAGGAGCAATCAACAGCTC
QY 531 CCCTCTGGCTTACAAACCCGAGATCGGAGTGTATAGTACCCCGGCTGGGCTCTT
Db 481 CCCTCTGGCTTACAAACCCGAGATCGGAGTGTATAGTACCCCGGCTGGGCTCTT
QY 591 CCTGTACTGTCTAG 603
Db 541 CCTGTACTGTCTAG 553
RESULT 13
AC127470 218485 bp DNA linear HTG 05:
LOCUS Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE,
DEFINITION ordered pieces.
AC127470
AC127470.4 GI:31415893
VERSION HTG: HTGS PHASE2; HTGS DRAFT.
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 218485)
AUTHORS Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjami
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Car
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurlburt,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-L
Legaspi,R., Maduro,Q.D., Maduro,V.B., Margulies,B.H., Masi
Maskeri,B., McDowell,J., Paquirigan,C., Pearson,R., Portno
Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., S
Sison,C., Stantrips,S., Thomas,J.W., Thomas,P.J., Teipour
Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 218485)
Green,E.D.
Direct Submission
Submitted (17-JUL-2002) NIH Intramural Sequencing Center,
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 218485)
Green,E.D.
Direct Submission
Submitted (05-JUN-2003) NIH Intramural Sequencing Center,
Government Circle, Gaithersburg, MD 20877, USA
On Jun 5, 2003 this sequence version replaced gi:26449071.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehhgrini.nih.gov
----- Project Information
Center project name: cms
Center clone name: 145D13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicat order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

[illegible]


```

/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="RP44-436K21"
/clone_lib="RP44"
1..27615
/notes="assembly_fragment
clone_end:SP6
vector_side:left"
27716..63581
/notes="assembly_fragment"
63682..79016
/notes="assembly_fragment"
79117..91018
/notes="assembly_fragment"
91119..94289
/notes="assembly_fragment"
94390..167164
/notes="assembly_fragment"
167265..180222
/notes="assembly_fragment
clone_end:T7
vector_side:right"

10.6%; Score 145; DB 2; Length 180222;
arity 100.0%; Pred.No. 1.2e-68;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|CCCCGGCGCTCCCCCTCCCGATCCCTCGGTCGCCGGATGCGGGGCGCGGTGAGG 84
|CCCCGGCGCTCCCCCTCCCGATCCCTCGGTCGCCGGATGCGGGGCGCGGTGAGG 95940
|CACAGCCCCCGCCCCCATGCGCGCCCGTCGAGCCAGAGCGGAGGGGCGCCGG 144
|CACAGCCCCCGCCCCCATGCGCGCCCGTCGAGCCAGAGCGGAGGGGCGCCGG 96000
|AGCGGGGACCGCCCTGCTGG 169
|AGCGGGGACCGCCCTGCTGG 96025

195 130254 bp DNA linear HTG 06-JUN-2003
s norvegicus clone RP31-258K6 strain Brown Norway, WORKING
SEQUENCE, 12 ordered pieces.
195
195.3 GI:31442440
HTGS_PHASE2; HTGS_DRAFT.
s norvegicus (Norway rat)
s norvegicus
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
ases 1 to 130254)
ellis;A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
sley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
pi.R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C.,
ri.B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
d.A., Reddi-Dugue,N., Schandler,K., Schueiler,M.G., Shah,K.,
., C., Stantripo,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
Comparative Sequencing Initiative
lished
ases 1 to 130254)
.E.D.
t Submission

```

Submitted (30-OCT-2002) NIH Intramural Sequencing Center,
GroveMont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 130254)
Green, B.D.
Direct Submission
Submitted (06-JUN-2003) NIH Intramural Sequencing Center,
GroveMont Circle, Gaithersburg, MD 20877, USA
On Jun 6, 2003 this sequence version replaced gi:27753660.
----- Genome Center -----
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information -----
Center project name: dcf
Center clone name: 258K06

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig, has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.930319
Consensus quality: 138169 bases at least Q40
Consensus quality: 138674 bases at least Q30
Consensus quality: 138972 bases at least Q20
Insert size: 150000; agarose-ftp
Insert size: 129154; sum-of-contigs
Quality coverage: 10.78x in Q20 bases; agarose-ftp
Quality coverage: 12.52x in Q20 bases; sum-of-contigs

```

- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 12 contigs. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that ha
- * provided by the submitter.

- * This sequence will be replaced.
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

*	1	10521: contig of 10521 bp in length
*	10522	10621: gap of unknown length
*	10622	13327: contig of 2706 bp in length
*	13328	13427: gap of unknown length
*	13428	28924: contig of 15497 bp in length
*	28925	29024: gap of unknown length
*	29025	39201: contig of 10177 bp in length
*	39202	39301: gap of unknown length
*	39302	41906: contig of 2605 bp in length
*	41907	42006: gap of unknown length
*	42007	70095: contig of 28089 bp in length
*	70096	70195: gap of unknown length
*	70196	77561: contig of 7366 bp in length
*	77562	77661: gap of unknown length
*	77662	94161: contig of 16500 bp in length
*	94162	94261: gap of unknown length
*	94262	97982: contig of 3721 bp in length
*	97983	98082: gap of unknown length
*	98083	105400: contig of 7318 bp in length
*	105401	105500: gap of unknown length
*	105501	127895: contig of 22355 bp in length
*	127895	127955: gap of unknown length
*	127956	130254: contig of 2299 bp in length.

Location/Qualifiers

```

1. 130254
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Brown Norway"
/db_xref="taxon:10116"
/clone="RP31-258K6"
/clone.lib="RP31"
1. 10521
/note="assembly_fragment"
clone_end:SP6
vector_side:left
10622..13327
/note="assembly_fragment"
13428..28924
/note="assembly_fragment"
29025..39201
/note="assembly_fragment"
39302..41906
/note="assembly_fragment"
42007..70095
/note="assembly_fragment"
70196..77561
/note="assembly_fragment"
77662..94161
/note="assembly_fragment"
94262..97982
/note="assembly_fragment"
98083..105400
/note="assembly_fragment"
105501..127855
/note="assembly_fragment"
117804..130254
/note="clone overlaps with GenBank Accession Number
AC134317 clone RP31-421B16 (center project name dce)"
127956..130254
/note="assembly_fragment
missing T7 clone end on 3' end of insert"

8.7%; Score 119; DB 2; Length 130254;
rity 100.0%; Pred. No. 3.9e-54; Indels 0; Gaps 0;
nservative 0; Mismatches 0;

TCGGGTCCCGGATGGCGGGCGGTGAGGACAGCCCGCCCGCCCATGGC 110
|||||
TCGGGTCCCGGATGGCGGGCGGTGAGGACAGCCCGCCCGCCCATGGC 64163
|||||

GTCGAGCCACAGCGGAGGGCGCGCGGGGGAGCGCGGACCGCCCTGCTGG 169
|||||
GTCGAGCCACAGCGGAGGGCGCGCGGGGGAGCGCGGACCGCCCTGCTGG 64222
|||||

5 165316 bp DNA linear HTG 19-NOV-2002
norvegicus clone CH230-320N23, WORKING DRAFT SEQUENCE, 5
ed pieces.

5.4 GI:25100662
GS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
norvegicus (Norway rat)
norvegicus
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

es 1 to 165316)
Marie., Metzker M.Lee., Abramson S., Adams C., Alder J.,
Allen H., Alsbrooks S., Amin A., Anguiano D.,
echi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Blair J., Blankenburg K., Blyth P., Brown M.,
N., Buhay C., Burch P., Burrell K., Calderon E.,

```

Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., C Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Sou Dava, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederic Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Ea Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, F., Fan, G Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Fo Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Gar Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Gue Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jac Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jol Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C. Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangu Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Na Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Park Nwackeleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Park Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Re Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Rui, S., S Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., S Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D. Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Staible, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Tayl Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uma Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., War Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoc Yu, F., Zhang, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H. Weinstein, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 165316)
Worley, K.C.
Direct Submission
Submitted (25-APR-2002) Human Genome Sequencing Center, Def of Molecular and Human Genetics, Baylor College of Medicine Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 165316)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Def of Molecular and Human Genetics, Baylor College of Medicine Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23616728.
The sequence in this assembly is a combination of BAC bases and whole genome shotgun sequencing reads assembled using / (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig is in the feature table below represents a scaffold in the ACI assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

b site: <http://www.hgsc.bcm.tmc.edu/>
 atact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 nter project name: GUXG
 nter clone name: CH230-320N23
 ----- Summary Statistics
 nsens quality: 159511 bases at least Q40
 nsens quality: 157321 bases at least Q30
 nsens quality: 158253 bases at least Q20
 nated insert size: 159662; sum-of-contigs estimation
 ality coverage: 7x in Q20 bases; sum-of-contigs estimation

 : Estimated insert size may differ from sequence length
 ee http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html.
 : This is a 'working draft' sequence. It currently
 lists of 5 contigs. The true order of the pieces
 ot known and their order in this sequence record is
 trary. Gaps between the contigs are represented as
 of N, but the exact sizes of the gaps are unknown.
 record will be updated with the finished sequence
 oon as it is available and the accession number will
 reserved.
 1 159511: contig of 159511 bp in length
 9512 159611: gap of unknown length
 9612 160669: contig of 1058 bp in length
 0670 160769: gap of unknown length
 0770 161864: contig of 1095 bp in length
 1865 161964: gap of unknown length
 1965 163701: contig of 1737 bp in length
 3702 163801: gap of unknown length
 3802 165316: contig of 1515 bp in length.
 Location/Qualifiers
 1. 165316
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-320N23"
 1. 2103
 /note="wgs_end_extension
 clone_end:17"
 8259. 9143
 /note="clone_boundary
 clone_end:17"
 site_
 end_sequence:B2171094"
 46341. 159014
 /note="clone_boundary
 clone_end:Sp6
 site:
 end_sequence:B2171096"

 8.7%; Score 119; DB 2; Length 165316;
 arity 100.0%; Pred. No. 3.9e-54; Indels 0; Gaps 0;
 onservative 0; Mismatches 0;
 CTCGGTCCCGGATGGGGGGCGGTGAGGAGGACAGCCCGCCCGCCCATGGC 110
 |||||
 CTCGGTCCCGGATGGGGGGCGGTGAGGAGGACAGCCCGCCCGCCCATGGC 120224
 |||||
 XGTCGAGCCAGAGCGGGGGCGCGGGGGGAGCGGGGACCGCCCTGCTGG 169
 |||||
 XGTCGAGCCAGAGCGGGGGCGCGGGGGGAGCGGGGACCGCCCTGCTGG 120165
 |||||
 i59 203083 bp DNA linear HTG 27-JUN-2001
 sculus chromosome 11 clone RP23-168P5, WORKING DRAFT
 VCF, 7 unordered pieces.
 i59
 i59.23 GI:14547768

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murin
 1 (bases 1 to 203083)
 Metzer, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,
 Dederich, D., Thomas, S., Okwuonu, G., Carlock, C., Garner, T.,
 Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Br
 Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z
 Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,
 Fernandez, C., Ferraguto, D., Forcum-Ranney, J., Gill, R.,
 Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogue
 Kovar, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J.,
 Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morri
 Neald, D., Nelson, A., Nguyen, R., Nguyen, N., Oguh, M., Parish,
 Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watling
 Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A.
 Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein,
 Worley, K. and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 203083)
 Worley, K.C.
 Direct Submission
 Submitted (31-MAY-2000) Human Genome Sequencing Center, De
 of Molecular and Human Genetics, Baylor College of Medicine
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 25, 2001 this sequence version replaced gi:12621364
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: MAFO
 Center clone name: RP23-168P5
 ----- Summary Statistics
 Sequencing vector: M13; L08821
 Chemistry: Dye-terminator Big Dye; 48% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 212648 bases at least Q40
 Consensus quality: 218902 bases at least Q30
 Consensus quality: 222384 bases at least Q20
 Estimated insert size: 210656; sum-of-contigs estimati
 Quality coverage: 0x in Q20 bases; agarose-fp estimati
 Quality coverage: 7.2x in Q20 bases; sum-of-contigs es

 * NOTE: Estimated insert size may differ from sequence ler
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_da
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 62152: contig of 62152 bp in length
 * 62153 62252: gap of unknown length
 * 62253 118772: contig of 56520 bp in length
 * 118773 118872: gap of unknown length
 * 118873 148924: contig of 30052 bp in length
 * 148925 149024: gap of unknown length
 * 149025 167231: contig of 18207 bp in length
 * 167232 167331: gap of unknown length
 * 167332 189007: contig of 22576 bp in length
 * 189008 190007: gap of unknown length
 * 190008 196537: contig of 6530 bp in length
 * 196538 196637: gap of unknown length
 * 196638 203083: contig of 6446 bp in length.

Location/Qualifiers
 1. .203083
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-168P5"

8.7% Score 119; DB 2; Length 203083;
 100.0%; Pred. No. 4e-54; Indels 0; Gaps 0;
 nservative 0; Mismatches 0;
 TCGGTCCTCCGGGATCGGGGGCGGTGAGGACGACAGCCCGCCCGCCATGGC 110
 TCGGTCCTCCGGGATCGGGGGCGGTGAGGACGACAGCCCGCCCGCCATGGC 51674
 TCGGAGCAGAGCGGAGGGGGCGCGGGGGGAGCGGACCGCCCTGTGG 169
 TCGGAGCAGAGCGGAGGGGGCGCGGGGGGAGCGGACCGCCCTGTGG 51615

223877 bp DNA linear HTG 10-MAY-2003
 norvegicus clone CH230-154B15, WORKING DRAFT SEQUENCE, 3
 ed pieces.

3.8 GI:30521223
 VGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

norvegicus (Norway rat)

Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

ies 1 to 223877)

.Marie, Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, H., Alsbrooks, S., Amin, A., Argulano, D.,
 Aebi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
 D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 is, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 and, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 , Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 lez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 orgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 ne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
 yez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 u.L., Jacob, L., Jiang, K., Johnson, B., Johnson, R., Jolivet, A.,
 W.S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 , Kraft, C.I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 ihwa, L., Louissegh, H., Lozano, R.J., Lu, X., Ma, J.,
 ari, M., Mahdardine, M., Mahmoud, M., Malloy, K., Mangum, A.,
 B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 ay, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 vjevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 vis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 lemech, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 nak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,
 r, F., Poidexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 X., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,

Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D
 Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soia, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Tay,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmi
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wa
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yo
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H
 Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 223877)

Worley, K.C.

Direct Submission

Submitted (06-NOV-2001) Human Genome Sequencing Center, De
 of Molecular and Human Genetics, Baylor College of Medicin
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 223877)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, De
 of Molecular and Human Genetics, Baylor College of Medicin
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:25008075
 The sequence in this assembly is a combination of BAC bases
 and whole genome shotgun sequencing reads assembled using
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig d
 in the feature table below represents a scaffold in the At
 assembly (a 'contig-scaffold'). Within each contig-scaffold
 individual sequence contigs are ordered and oriented, and
 by sized gaps filled with Ns to the estimated size. The se
 may extend beyond the ends of the clone and there may be s
 contigs within a contig-scaffold that consist entirely of
 genome shotgun sequence reads. Both end sequences and whol
 shotgun sequence only contigs will be indicated in the fea
 table.

Center: Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GIOK

Center Clone name: CH230-154B15

Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 214785 bases at least Q40

Consensus quality: 216908 bases at least Q30

Consensus quality: 218593 bases at least Q20

Estimated insert size: 227169; sum-of-contigs estimati

Quality coverage: 7x in Q20 bases; sum-of-contigs esti

NOTE: Estimated insert size may differ from sequence len

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_da

NOTE: This is a 'working draft' sequence. It currently

consists of 3 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence.

as soon as it is available and the accession number will

be preserved.

1 221327: contig of 221327 bp in length

221328 221427: gap of unknown length

221428 222652: contig of 1225 bp in length

222653 222752: gap of unknown length

222753 223877: contig of 1125 bp in length.

Location/Qualifiers

1. 223877

/organism="Rattus norvegicus"

source


```

/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-154B15"
1..1822
/note="wgs contig"
complement(217607..218056)
/note="clone boundary"
clone_end:77
site:EcORI
end_sequence:RWBB0087B"

      8.7%  Score 119;  DB 2;  Length 223877;
arity 100.0%;  Pred. No. 4e-54;
conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

CTCGGTCCTCCGGATGGGGGGGGCGGTGAGGAGGACAGAGCCCGCCCGCCCATGGC 110
|||||
CTCGGTCCTCCGGATGGGGGGGGCGGTGAGGAGGACAGAGCCCGCCCGCCCATGGC 139398
|||||

CTCGGAGCCAGAGCGGAGGGGGCGCCGGGGGAGCGCGGACACCGCCCTGCTGG 169
|||||
CTCGGAGCCAGAGCGGAGGGGGCGCCGGGGGAGCGCGGACACCGCCCTGCTGG 139457
|||||

63          225077 bp      DNA      linear      HTG 10-MAY-2003
norvegicus clone CH230-46E21, WORKING DRAFT SEQUENCE.

63.3  GI:30521905
HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
norvegicus (Norway rat)
norvegicus
ota; Metazoa; Chordata; Vertebrata; Euteleostomi;
lia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
1. 225077)
D.Marie, Metzker M.Lee, Abramson S., Adams C., Alder J.,
C.Allen, H., Alabrooks S., Amin A., Anguiano D.,
bechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
in D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
to K., Blair J., Blankenburg K., Blyth P., Brown M.,
C.N., Buhal J., Burch P., Burrell K., Calderon E.,
as V., Carter K., Cavazos I., Ceasar H., Center A.,
o J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
lad C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
a M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
to O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,
r H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
ndez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
r C.M., Gabies A., Ganta R., Garcia A., Garner T., Garza M.,
George E., Geer K., Gill R., Grady M., Guerra W., Guevara M.,
atne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
y Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
ndez R., Hines S., Hladun S.L., Hodgson A., Hoques M.,
ns B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
on L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
thy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
Liu W., Liu Y., London P., Longacre S., Lopez J.,
shewa L., Louiseged H., Lozada R.J., Lu X., Ma J.,
hwari M., Mahindartine M., Mahmoud M., Malloy K., Mangum A.,
m B., Mapua P., Martin K., Martin R., Martinez E.,
ney S., McLeod M.P., McNeill T.Z., Meenen E.,
aviljevic A., Miner G., Minja E., Montemayor J., Moore S.,
n M., Morris K., Morris S., Munidasa M., Murphy M., Nair L.,
elish C., Neal D., Newton N., Nguyen N., Norris S.,
eleneh O., Okwuonu G., Olarnpunsagoon A., Pal S., Parks K.,
rnak S., Paul H., Perez A., Perez L., Pfannkoch C.,
er F., Poindexter A., Popovic D., Primus E., Pu L.-L.,
M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Rat Genome Sequencing Consortium.
Rat Genome Sequencing Consortium.
Submitted (06-NOV-2002) Human Genome Sequencing Center, De
of Molecular and Human Genetics, Baylor College of Medicin
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 225077)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, De
of Molecular and Human Genetics, Baylor College of Medicin
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC base
and whole genome shotgun sequencing reads assembled using
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig (
in the feature table below represents a scaffold in the A
assembly (a 'contig-scaffold'). Within each contig-scaffo
individual sequence contigs are ordered and oriented, and
by sized gaps filled with Ns to the estimated size. The s
may extend beyond the ends of the clone and there may be
contigs within a contig-scaffold that consist entirely of
genome shotgun sequence reads. Both end sequences and who
shotgun sequence only contigs will be indicated in the fe
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHDQ
Center clone name: CH230-46E21
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 214599 bases at least Q40
Consensus quality: 216978 bases at least Q30
Consensus quality: 218480 bases at least Q20
Estimated insert size: 225046; sum-of-contigs estimat
Quality coverage: 8x in Q20 bases; sum-of-contigs est

* NOTE: Estimated insert size may differ from sequence le
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_d
* NOTE: This is a 'working draft' sequence. It currently
consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that ha
provided by the submitter.
* This sequence will be replaced
* By the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 225077: contig of 225077 bp in length.
Location/Qualifiers
1..225077
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
FEATURES
source

```

```

/db_xref="taxon:10116"
/clone="CH230-46E21"
1. .1402
/note="wgs_contig"

      8.7%; Score 119; DB 2; Length 225077;
urity 100.0%; Pred. NO. 4e-54;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

TTCGGTCCCGGATGGGGGGCGGTGAGGACGACACAGCCCGCCCGCCCATGGC 110
TTCGGTCCCGGATGGGGGGCGGTGAGGACGACACAGCCCGCCCGCCCATGGC 199986

GTGCGAGCAGAGCGAGGGGGCGCGGGGGAGCCGGGACCCCGCTGCTGG 169
GTGCGAGCAGAGCGAGGGGGCGCGGGGGAGCCGGGAGCCGGGACCCCGCTGCTGG 199927

07      234182 bp DNA linear ROD 17-NOV-2001
DNA sequence from clone RP23-422L16 on chromosome 11,
ie sequence.
07.5 GI:17017790
iculus (house mouse)
ta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
a; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
es 1 to 234182)
A.
Submission
ed (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
ighshire, CB10 1SA, UK. E-mail enquiries:
ysanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
20, 2001 this sequence version replaced gi:16605765.
sequence assembly data is compared from overlapping clones.
ifferences are found these are annotated as variations
on with a note of the overlapping clone name. Note that the
on annotation may not be found in the sequence submission
onding to the overlapping clone, as we submit sequences with
small overlap as described above.
sequence was finished as follows unless otherwise noted: all
s were either double-stranded or sequenced with an alternate
ry or covered by high quality data (i.e., phred quality >=
1 attempt was made to resolve all sequencing problems, such
pressions and repeats; all regions were covered by at least
amid subclone or more than one M13 subclone; and the
ly was confirmed by restriction digest. The following
iations are used to associate primary accession numbers given
feature table with their source databases: Em, EMBL; Sw,
ROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
se can be found at
/www.sanger.ac.uk/Projects/C_elegans/wormpep/RP23-422L16 is
ne RPCI-23 Mouse PAC Library
acted by the group of Pieter de Jong.
rther details see http://www.chori.org/bacpac/home.htm
: pBAce3.6
Location/Qualifiers
1. .234182
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_lib="RPCI-23"
complement(84050..84131)
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."

```

```

Query Match      8.7%; Score 119; DB 10; Length 234182;
Best Local Similarity 100.0%; Pred. NO. 4.1e-54;
Matches 119; Conservative 0; Mismatches 0; Indels 0; G
QY      51 ATCCCTCCGGTCCCGGATGGGGGGCGGTGAGGACGACACAGCCCGCCCGCCCATGGC
DB      75038 ATCCCTCCGGTCCCGGATGGGGGGCGGTGAGGACGACACAGCCCGCCCGCCCATGGC
QY      111 CGCCCTCCGAGCAGAGCGGAGGGGGCGCGGGGGAGCCCGCTGCTGG
DB      74978 CGCCCTCCGAGCAGAGCGGAGGGGGCGCGGGGGAGCCCGCTGCTGG

RESULT 22
AC126237      212093 bp DNA linear HTG 06
LOCUS
DEFINITION
Canis familiaris clone RP81-414022, WORKING DRAFT SEQUENCE
ordered pieces.
AC126237
AC126237.5 GI:31442445
VERSION
HTG; HTGS PHASE2; HTGS_DRAFT.
KEYWORDS
Canis familiaris (dog)
SOURCE
Canis familiaris
ORGANISM
Canis familiaris
REFERENCE
1 (bases 1 to 212093)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele-
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjami-
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Car-
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurlb,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-L
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masi
Maskeri,B., McDowell,J., Paquirigan,C., Pearson,R., Porto
Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., S
Sison,C., Stantropop,S., Thomas,J.W., Thomas,P.J., Tsipour
Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 212093)
Green,E.D.
Direct Submission
Submitted (04-JUL-2002) NIH Intramural Sequencing Center,
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 212093)
Green,E.D.
Direct Submission
Submitted (06-JUN-2003) NIH Intramural Sequencing Center,
Grovemont Circle, Gaithersburg, MD 20877, USA
On Jun 6, 2003 this sequence version replaced gi:27476131.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@hgri.nih.gov
----- Project Information
Center project name: CWG
Center clone name: 414022

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicat order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads

```
misc_feature      |note="assembly_fragment"  
142963.352284    |note="assembly_fragment"  
misc_feature      |note="assembly_fragment"  
152685.155868    |note="assembly_fragment"  
misc_feature      |note="assembly_fragment"  
155769.176357    |note="assembly_fragment"  
misc_feature      |note="assembly_fragment"  
176658.184091    |note="assembly_fragment"  
misc_feature      |note="assembly_fragment"  
184192.195037    |note="assembly_fragment"  
misc_feature      |note="assembly_fragment"  
195138.200818    |note="assembly_fragment"  
misc_feature      |note="assembly_fragment"  
200919.204476    |note="assembly_fragment"  
misc_feature      |note="assembly_fragment"  
204577.212093    |note="assembly_fragment"  
clone_end:T7  
vector_side:right"
```

ORIGIN

Query Match 4.7%; Score 65; DB 2; Length 212093;
Best Local Similarity 100.0%; Pred. No. 6e-24;
Matches 65; Conservative 0; Mismatches 0; Indels 0;

QY 85 CAGGCACAGCCCCCCCCCATGGCCGCGCCCGTCGGAGCCAGAGCGGAGGGGGG
|||
176827 CAGGCACAGTCCCCCGCCCCCATGGCCGCGCCCGTCGGAGCCAGAGCGGAGGGGGG

145 GGGGA 149
Ov

Db 176887 GGGGA 176891

RESULT 23

AF030100	LOCUS	AF030100	1239 bp	mrna	linear	ROD 3
	DEFINITION	Mus musculus	TWEAK	mrna.	complete	cds.

AF030100.2 GI:33348855
ACCESSION VERSION

KEYWORDS
SOURCE

ORGANISM

100

AUTHORS

TITLE

JOURNAL

PITRMED
MEDLINE

REFERENCE

TITLE	COMMITTEE
1. The Role of the State in the Development of the Economy	1. The Role of the State in the Development of the Economy
2. The Role of the State in the Development of the Economy	2. The Role of the State in the Development of the Economy
3. The Role of the State in the Development of the Economy	3. The Role of the State in the Development of the Economy
4. The Role of the State in the Development of the Economy	4. The Role of the State in the Development of the Economy
5. The Role of the State in the Development of the Economy	5. The Role of the State in the Development of the Economy
6. The Role of the State in the Development of the Economy	6. The Role of the State in the Development of the Economy
7. The Role of the State in the Development of the Economy	7. The Role of the State in the Development of the Economy
8. The Role of the State in the Development of the Economy	8. The Role of the State in the Development of the Economy
9. The Role of the State in the Development of the Economy	9. The Role of the State in the Development of the Economy
10. The Role of the State in the Development of the Economy	10. The Role of the State in the Development of the Economy

AUTHORS

JOURNAL

REMARK

COMMENT
FEATURES

2003

33


```

/note="assembly_fragment"
126979..148555
/note="clone overlaps with GenBank Accession Number
AC134961 clone RP42-406J16 (center project name djv)"
126979..129099
/note="assembly_fragment"
129200..135799
/note="assembly_fragment"
135900..145102
/note="assembly_fragment"
145203..148555
/note="assembly_fragment
clone_end:77
vector_side:right"

4.2%; Score 58; DB 2; Length 148555;
larity 100.0%; Pred. No. 4.7e-20;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ATGACCTTGTGATGAGGGAGGCTGTCTACCTGAAGCTGGACTGCTGTGGATG 658
|||||
ATGACCTTGTGATGAGGGAGGCTGTCTACCTGAAGCTGGACTGCTGTGGATG 60929

925 176258 bp DNA linear HTG 06-JUN-2003
familiaris clone RP81-332E11, WORKING DRAFT SEQUENCE, 12
ed pieces.
925
925.6 GI:31442444
HTGS PHASE2; HTGS_DRAFT.
familiaris (dog)
familiaris
yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
lia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
ases 1 to 176258)
sley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
pi.R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C.,
ri.B., McDowell,J., Paquinigan,C., Pearson,R., Portnoy,M.E.,
d.A., Reddix-Dugae,N., Schandler,K., Schueler,M.G., Shah,X.,
C., Stantripoop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
Comparative Sequencing Initiative
lished
ases 1 to 176258)
t Submission
tted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
mont Circle, Gaithersburg, MD 20877, USA
ases 1 to 176258)
t Submission
tted (06-JUN-2003) NIH Intramural Sequencing Center, 8717
mont Circle, Gaithersburg, MD 20877, USA
n 6, 2003 this sequence version replaced gi:28209436.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoonhgri.nih.gov
----- Project Information
Center project name: cwp
Center clone name: 332E11

```

sequence data in this record represents an 'enhanced' on of a Phase 2 submission. Specifically, the indicated and orientation of each sequence contig has been

established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g. human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 173760 bases at least Q40
Consensus quality: 174423 bases at least Q30
Consensus quality: 174916 bases at least Q20
Insert size: 152000; agarose-fp
Insert size: 175158; sum-of-contigs
Quality coverage: 17.46x in Q20 bases; agarose-fp
Quality coverage: 15.15x in Q20 bases; sum-of-contig

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

```

1 11425: contig of 11425 bp in length
* 11426 11525: gap of unknown length
* 11526 27554: contig of 16029 bp in length
* 27555 29974: gap of unknown length
* 29975 30075: contig of 2320 bp in length
* 30076 36241: gap of unknown length
* 36242 36342: contig of 6167 bp in length
* 36343 77002: contig of 40661 bp in length
* 77003 77103: gap of unknown length
* 77104 117018: contig of 39916 bp in length
* 117019 117119: gap of unknown length
* 117120 119041: contig of 1923 bp in length
* 119042 119141: gap of unknown length
* 119142 158488: contig of 39247 bp in length
* 158489 168033: contig of 9545 bp in length
* 168034 168133: gap of unknown length
* 168134 170716: contig of 2583 bp in length
* 170717 170816: gap of unknown length
* 170817 174429: contig of 3613 bp in length
* 174430 174529: gap of unknown length
* 174530 176258: contig of 1729 bp in length.

```

FEATURES

source

```

1..176258
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
/clone="RP81-332E11"
/clone_lib="RP81"

```

misc_feature

```

1..88033
/note="clone overlaps with GenBank Accession Num
AC126237 clone RP81-414022 (center project name
1..11425
/note="assembly_fragment

```

misc_feature

```

clone_end:896
vector_side:left
11526..17554
/note="assembly_fragment"
27655..29974
/note="assembly_fragment"
30075..36241
/note="assembly_fragment"

```

misc_feature

```

11526..17554
/note="assembly_fragment"
27655..29974
/note="assembly_fragment"
30075..36241
/note="assembly_fragment"

```

misc_feature

```

11526..17554
/note="assembly_fragment"
27655..29974
/note="assembly_fragment"
30075..36241
/note="assembly_fragment"

```

misc_feature

```

11526..17554
/note="assembly_fragment"
27655..29974
/note="assembly_fragment"
30075..36241
/note="assembly_fragment"

```

```

36342..77002
/note="assembly_fragment"
77103..117018
/note="assembly_fragment"
/note="assembly_fragment"
117119..119041
/note="assembly_fragment"
119142..158388
/note="assembly_fragment"
158489..168033
/note="assembly_fragment"
168134..170716
/note="assembly_fragment"
170817..174429
/note="assembly_fragment"
174530..176258
/note="assembly_fragment
vector_side:right"

4.2k; Score 58; DB 2; Length 176258;
arity 100.0%; Pred. No. 4.8e-20; Indels 0; Gaps 0;
conservative 0; Mismatches 0;

IGCATTGTGATGAGGGAAGGCTGTCTACCTGAAGCTGGAGCTGCTGGTGATG 658
|||||
IGCATTGTGATGAGGGAAGGCTGTCTACCTGAAGCTGGAGCTGCTGGTGATG 59938

39          149736 bp      DNA      linear      HTG 06-JUN-2003
catus clone RP86-474H17, WORKING DRAFT SEQUENCE, 12 ordered

39          39.5 GI:31442446
IGS PHASE2; HTGS_DRAFT.
catus (cat)

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Euteleia; Carnivora; Fissipedia; Felidae; Felis.
ses 1 to 149736
llis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
ley,R.W., Bouffard,G.G., Brinkley,C., Brooke,S., Cariaga,K.,
Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
i,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
i,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.B.,
A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
Comparative Sequencing Initiative
ished
ses 1 to 149736)
E.D.
Submission
ted (04-JUL-2002) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
ses 1 to 149736)
E.D.
Submission
ted (06-JUN-2003) NIH Intramural Sequencing Center, 8717
ont Circle, Gaithersburg, MD 20877, USA
.6, 2003 this sequence version replaced gi:26801264.
----- Genome Center
'enter: NIH Intramural Sequencing Center
'eb site: http://www.nisc.nih.gov
'ontact: nisc@ohgri.nih.gov
----- Project Information
'enter project name: daa
'enter clone name: 474H17

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicat order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 148041 bases at least Q40

Consensus quality: 148385 bases at least Q30

Consensus quality: 148558 bases at least Q20

Insert size: 126000; agarose-fp

Quality coverage: 15.56x in Q20 bases; agarose-fp

Quality coverage: 13.19x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

1	483:	contig of 483 bp in length
484	583:	gap of unknown length
584	19266:	contig of 18683 bp in length
19267	19366:	gap of unknown length
19367	21971:	contig of 2605 bp in length
21972	22071:	gap of unknown length
22072	24316:	contig of 2245 bp in length
24317	24416:	gap of unknown length
24417	36979:	contig of 12563 bp in length
36980	37079:	gap of unknown length
37080	42619:	contig of 5540 bp in length
42620	45464:	contig of 3745 bp in length
45465	46564:	gap of unknown length
46565	61551:	contig of 14987 bp in length
61552	61651:	gap of unknown length
61652	99225:	contig of 37574 bp in length
99226	99326:	gap of unknown length
99326	110459:	contig of 11134 bp in length
110460	110553:	gap of unknown length
110560	147587:	contig of 37028 bp in length
147588	147687:	gap of unknown length
147688	149736:	contig of 2049 bp in length.

Location/Qualifiers

1..149736

/organism="Felis catus"

/mol_type="genomic DNA"

/db_xref="taxon:9685"

/clone="RP86-474H17"

/clone_lib="RP86"

1..483

/note="assembly_fragment"

clone_end:T7

vector_side:left

584..19266

/note="assembly_fragment"

19367..21971

/note="assembly_fragment"

22072..24316

/note="assembly_fragment"

24317..36979

/note="assembly_fragment"

36980..42619

/note="assembly_fragment"

42620..45464

/note="assembly_fragment"

45465..61551

/note="assembly_fragment"

61552..99225

/note="assembly_fragment"

99226..110459

/note="assembly_fragment"

110460..110553

/note="assembly_fragment"

110554..147587

/note="assembly_fragment"

147588..147687

/note="assembly_fragment"

147688..149736

/note="assembly_fragment"

FEATURES

Source

misc_feature

misc_feature

misc_feature

misc_feature

/note="assembly_fragment"
24417. .36979
/note="assembly_fragment"
37080. .42619
/note="assembly_fragment"
42720. .46464
/note="assembly_fragment"
46565. .61551
/note="assembly_fragment"
61652. .99225
/note="assembly_fragment"
99326. .110459
/note="assembly_fragment"
110560. .147587
/note="assembly_fragment"
147688. .149736
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

3.9%; Score 53; DB 2; Length 149736;
Identity 100.0%; Pred. No. 3e-17; 0; Indels 0; Gaps 0;
Conservative 0; Mismatches 0;

CCAGGTCCTCTCGGATCCGACCCCTCCCTCGGCGCCATCTCAAGGC 812
|||||
CCAGGTCCTCTCGGATCCGACCCCTCCCTCGGCGCCATCTCAAGGC 43768

395 ice 74 from Patent WO0153486. linear PAT 30-AUG-2001
395
395.1 GI:15391196

atic construct
tic construct
ical sequences.

razi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
1,K.J., Marsters,S.A., Pan,J., Pitti,R.W., Roy,M.A., Smith,V.,
D.M., Watanabe,C.K. and Wood,W.I.
sitions and methods for the treatment of tumour
-: WO 0153486-A 74 26-JUL-2001;
tech, Inc. (US)

Location/Qualifiers
1. .50
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetic Oligonucleotide Probe."

3.6%; Score 50; DB 6; Length 50;
Identity 100.0%; Pred. No. 7e-16;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GCCTCTCGGTACACCGCAGATCGGGAGTTTATAGTCACCCGG 576
|||||
GCCTCTCGGTACACCGCAGATCGGGAGTTTATAGTCACCCGG 50

954 ligand polypeptide. linear PAT 27-AUG-2002
954
954.1 GI:22636564
01522584-A/3.
ntified
ntified

unclassified.

1 (bases 1 to 50)
Ashkenazi,A.J., Marsters,S.A. and Pitti,R.
Apo-3 ligand polypeptide
Patent: JP 2001522584-A 3 20-NOV-2001;
GENENTECH INC

OS Unknown
PN JP 2001522584-A/3
PD 20-NOV-2001
PF 09-OCT-1998 JP 2000516042
PR 10-OCT-1997 US 60/062037,17-DEC-1997 US 60/0691
AVI J ASHKENAZI, SCOT A MARSTERS, ROBERT PITTI
PC C12N15/09,A61K38/00,C07K14/705,C07K16/24,C12N15/00,A6
CC Description of Unknown Organism:Unknown
FH Key Location/Qualifiers
FT source 1. .50
/organism="Unknown".

FEATURES
source

1. .50
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 3.6%; Score 50; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 7e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0;

QY 527 CCAGCCCTCTCGGTACACCGCAGATCGGGAGTTTATAGTCACCCGG 576
|||||
DB 1 CCAGCCCTCTCGGTACACCGCAGATCGGGAGTTTATAGTCACCCGG 50

RESULT 29

AC129071/c
LOCUS

DEFINITION Pan troglodytes clone RP43-149M23, WORKING DRAFT SEQUENCE
ordered pieces.

ACCESSION

AC129071

AC129071.2 GI:26449072

VERSION

HTG; HTGS PHASE2; HTGS DRAFT.

KEYWORDS

SOURCE

ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE

AUTHORS

1 (bases 1 to 163542)

Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S

Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C.,

Benjamin,B., Coleman,B., Engle,J., Granite,S., Guan,X., Gu

Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karl

Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,

Margulies,E.H., Masello,C., Maskeri,B., McDowell,J.,

Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,

Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,

Standtripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., V

Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

UNPUBLISHED

2 (bases 1 to 163542)

Green,E.D.

Direct Submission

Submitted (25-JUL-2002) NIH Intramural Sequencing Center,

Groveton Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 163542)

Green,E.D.

Direct Submission

Submitted (11-DEC-2002) NIH Intramural Sequencing Center,

Groveton Circle, Gaithersburg, MD 20877, USA

On Dec 11, 2002 this sequence version replaced gi:2195501

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

contact: misc.zooenhgri.nih.gov
----- Project Information
enter project name: cmt
enter clone name: 149M23

quence data in this record represents an 'enhanced'
1 of a Phase 2 submission. Specifically, the indicated
and orientation of each sequence contig has been
ished using one or more of the following: read-pair
com individual subclones, overlaps with neighboring
alignment with available reference sequence (e.g.,
, and/or confirmation by PCR testing. In addition,
quence assembly is based on at least 8x average
ye in Q20 bases and has been reviewed to rule out
niassemblies, the low-quality ends of sequence
s have been trimmed away, and each base is associated
Phrap-derived quality score.

----- Summary Statistics
sequencing vector: plasmid; n/a; 100% of reads
chemistry: Dye-terminator Big Dye; 100% of reads
assembly program: Phrap; version 0.990319
consensus quality: 159675 bases at least Q40
consensus quality: 160848 bases at least Q30
insert size: 185000; agarose-fp
insert size: 162442; sum-of-contigs
ality coverage: 7.43x in Q20 bases; agarose-fp
ality coverage: 8.47x in Q20 bases; sum-of-contigs

: This is a 'working draft' sequence. It currently
lists of 12 contigs. Gaps between the contigs
represented as runs of N. The order of the pieces
elieved to be correct as given, however the sizes
he gaps between them are based on estimates that have
ided by the submittor.

sequence will be replaced
he finished sequence as soon as it is available and
accession number will be preserved.

1 10517: contig of 10517 bp in length
0518 10617: gap of unknown length
0618 25192: contig of 14575 bp in length
5193 25292: gap of unknown length
5293 27192: contig of 1900 bp in length
7193 27292: gap of unknown length
7293 29759: contig of 2467 bp in length
9760 29859: gap of unknown length
9860 57864: contig of 28005 bp in length
7865 57964: gap of unknown length
7965 61287: contig of 3323 bp in length
1288 61387: gap of unknown length
1388 84039: contig of 22652 bp in length
4040 84139: gap of unknown length
4140 114016: contig of 29877 bp in length
4017 114116: gap of unknown length
4117 116474: contig of 2358 bp in length
6475 116574: gap of unknown length
6575 121369: contig of 4795 bp in length
1370 121469: gap of unknown length
1470 154199: contig of 32730 bp in length
4200 154299: gap of unknown length
4300 163542: contig of 9243 bp in length.

Location/Qualifiers
1..163542
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-149M23"
/clone_lib="RP43"
1..10517
/note="assembly_fragment
missing T7 clone end on 5' end of insert"
10618..25192
/note="assembly_fragment"

misc_feature 25293..27192
/note="assembly_fragment"
misc_feature 27293..29759
/note="assembly_fragment"
misc_feature 29860..57864
/note="assembly_fragment"
misc_feature 57965..61287
/note="assembly_fragment"
misc_feature 61388..84039
/note="assembly_fragment"
misc_feature 84140..114016
/note="assembly_fragment"
misc_feature 114117..116474
/note="assembly_fragment"
misc_feature 116575..121369
/note="assembly_fragment"
misc_feature 121470..154199
/note="assembly_fragment"
misc_feature 154300..163542
/note="assembly_fragment
clone end:SP6
vector_side:right"

ORIGIN

Query Match 3.5%; Score 48; DB 2; Length 163542;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; G
QY 1312 CCTGTGGATTTTAAACAGATATTTTATTATTATTATTGTGACAAA 1359
Db 163542 CCTGTGGATTTTAAACAGATATTTTATTATTATTATTGTGACAAA 163495

RESULT 30

BD062757
LOCUS BD062757 1168 bp DNA linear PAT 27
DEFINITION A tumor necrosis factor related ligand.
ACCESSION BD062757
VERSION BD062757.1 GI:22608360
KEYWORDS JP 2001505407-A/1.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 1168)
AUTHORS Chicheportiche,Y. and Browning,J.L.
TITLE A tumor necrosis factor related ligand
JOURNAL Patent: JP 2001505407-A 1 24-APR-2001;
BIOGEN INC,THE FACULTY OF MEDICINE OF THE UNIVERSITY OF GE
COMMENT OS TNF family related protein
PN JP 2001505407-A/1
PD 24-APR-2001
PF 07-AUG-1997 JP 1998508239
PR 07-AUG-1996 US 60/023541,18-OCT-1996 US 60/0285
18-MAR-1997 US 60/040820
PI YVES CHICHEPORTICHE,JEFFREY L.BROWNING
PC C12N15/28,C07K14/525,G01N33/68,C07K16/24,C12N15/11,A6
PC C12N5/10,
PC A61K39/395,A61K38/19,C07K14/705,C12N15/12
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1168
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES

source
Query Match 3.4%; Score 46; DB 6; Length 1168;
Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; C
ORIGIN
Query Match 3.4%; Score 46; DB 6; Length 1168;
Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; C

TGGGCTCTACTGCTACTGCTACTGCTAGGTCACCTTTGATGAGG 620
 TGGGCTCTACTGCTACTGCTACTGCTAGGTCACCTTTGATGAGG 444

09 234801 bp DNA linear HTG 13-NOV-2002
 norvegicus clone CH230-212018, *** SEQUENCING IN PROGRESS
 unordered pieces.

09.6 GI:24941765
 HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 norvegicus (Norway rat)

norvegicus
 Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 ia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

uses 1 to 234801

D.Marie, Metzker, M.Lee, Abramson, S., Adams, C., Alder, J.,
 C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 ebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 n.D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 o.K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 as, V., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 land, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 i.M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 io.O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 a., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 a., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 idex, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 c.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 eorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 atne, P., Haaland, P., Hawes, A., Henderson, N., Hamilton, K.,
 f, Y., Havlak, P., Hawes, A., Henderson, N., Hamilton, K.,
 adez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 as, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 on, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 chy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 C., Kraft, C.I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 u, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 suhewa, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J.,
 hware, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 n, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 ney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 avljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 n, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 rvis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 elemeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 rnak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 er, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 m., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 y, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 rs, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 y, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 le, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 i, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 ms, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
 t, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 hauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Stock, G. and Gibbs, R.A.

Submitted

uses 1 to 234801

y.K.C.

Submitted

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (15-APR-2002) Human Genome Sequencing Center, De
 of Molecular and Human Genetics, Baylor College of Medicin
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 234801)
 Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-NOV-2002) Human Genome Sequencing Center, De
 of Molecular and Human Genetics, Baylor College of Medicin
 Baylor Plaza, Houston, TX 77030, USA

On Nov 13, 2002 this sequence version replaced gi:23267374
 and whole genome shotgun sequencing reads assembled using
 in the feature table below represents a scaffold in the At
 assembly (a 'contig-scaffold'). Within each contig-scaffo
 individual sequence contigs are ordered and oriented, and
 by sized gaps filled with Ns to the estimated size. The s
 may extend beyond the ends of the clone and there may be
 contigs within a contig-scaffold that consist entirely of
 genome shotgun sequence reads. Both end sequences and who
 shotgun sequence only contigs will be indicated in the fea
 table.

Center: Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GUTA

Center clone name: CH230-212018

Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 227577 bases at least Q40

Consensus quality: 230382 bases at least Q30

Consensus quality: 231817 bases at least Q20

Estimated insert size: 239036; sum-of-contigs estimat

Quality coverage: 5x in Q20 bases; sum-of-contigs est

NOTE: Estimated insert size may differ from sequence le

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft

NOTE: This is a 'working draft' sequence. It currently

consists of 5 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number wil

be preserved.

1 162694: contig of 162694 bp in length

162695 162794: gap of unknown length

162795 225625: contig of 62831 bp in length

225626 225725: gap of unknown length

225726 231241: contig of 5516 bp in length

231242 231341: gap of unknown length

231342 232640: contig of 1299 bp in length

232641 232740: gap of unknown length

232741 234801: contig of 2061 bp in length.

Location/Qualifiers

1. 234801

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-212018"

complement(160419..161185)

/notes="clone_boundary"

clone_end:T7

site:

end_sequence:BH264302"

225726..227372

/notes="wgs_end_extension"

clone_end:T7"

misc_feature

misc_feature

225726..227372

/notes="wgs_end_extension"

clone_end:T7"

ORIGIN

Db	67	ATGTCATTGTTAGACTTTGAAATTC	42
RESULT 34			
AX522345			
LOCUS	AX522345	145 bp	DNA linear PAT 24
DEFINITION	Sequence 15 from Patent WO02064731.		
ACCESSION	AX522345		
VERSION	AX522345.1	GI:24411299	
KEYWORDS			
SOURCE	Human sapiens (human)		
ORGANISM	Human sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1. Telerman, A., Anson, R., Tuijnder, M. and Susini, L.		
	Sequences involved in phenomena of tumour suppression, tum		
	reversion, apoptosis and/or virus resistance and their use		
	medicines		
JOURNAL	Patent: WO 02064731-A 15 22-AUG-2002;		
FEATURES	Molecular Engines Laboratories (FR)		
source	Location/Qualifiers		
	1. .145		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	1.9%;	Score 26;	DB 6; Length 145;
Best Local Similarity	100.0%;	Pred. No. 0.02;	
Matches	26;	Conservative 0;	Mismatches 0; Indels 0; G 0;
QY	1	ATGTCATTGTTAGACTTTGAAATTC	26
Db	74	ATGTCATTGTTAGACTTTGAAATTC	99
RESULT 35			
AX522465			
LOCUS	AX522465	145 bp	DNA linear PAT 24
DEFINITION	Sequence 135 from Patent WO02064731.		
ACCESSION	AX522465		
VERSION	AX522465.1	GI:24411419	
KEYWORDS			
SOURCE	Human sapiens (human)		
ORGANISM	Human sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1. Telerman, A., Anson, R., Tuijnder, M. and Susini, L.		
	Sequences involved in phenomena of tumour suppression, tum		
	reversion, apoptosis and/or virus resistance and their use		
	medicines		
JOURNAL	Patent: WO 02064731-A 135 22-AUG-2002;		
FEATURES	Molecular Engines Laboratories (FR)		
source	Location/Qualifiers		
	1. .145		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	1.9%;	Score 26;	DB 6; Length 145;
Best Local Similarity	100.0%;	Pred. No. 0.02;	
Matches	26;	Conservative 0;	Mismatches 0; Indels 0; G 0;
QY	1	ATGTCATTGTTAGACTTTGAAATTC	26
Db	74	ATGTCATTGTTAGACTTTGAAATTC	99
RESULT 36			

ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele-	
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo;	
1 (bases 1 to 353)	
REFERENCE	
AUTHORS	
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,	
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schi-	
Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bha-	
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsie-	
Diachenko,I., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,	
Stapleton,M., Soares,M.B., Donald,M.F., Casavant,T.L.,	
Scheetz,T.B., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,	
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peteri-	
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,	
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,	
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,	
Villaalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.	
Fahney,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,	
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,	
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,	
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myl-	
Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smailus,D.I.	
Schmurch,A., Schein,J.E., Jones,S.J. and Marra,M.A.	
Generation and initial analysis of more than 15,000 full-	
human and mouse cDNA sequences	
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
22388257	
12477932	
2 (bases 1 to 353)	
REFERENCE	
AUTHORS	
Strausberg,R.	
Direct Submission	
Submitted (17-NOV-2003) National Institutes of Health, Ma-	
Gene Collection (MGC), Cancer Genomics Office, National C-	
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 2089	
USA	
NIH-MGC Project URL: http://mgc.nci.nih.gov	
Contact: MGC help desk	
Email: cgapps-remail.nih.gov	
Tissue Procurement: ATCC/DCTD/PTP	
cDNA Library Preparation: CLONTECH Laboratories, Inc.	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: Genome Sequence Centre,	
BC Cancer Agency, Vancouver, BC, Canada	
info@cgsc.bc.ca	
Steven Jones, Jennifer Amano, Ian Bosdet, Yaron Butterfie	
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ra-	
Letitia Hsiao, Martin Krzyzinski, Reta Kutche, Oliver L	
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy,	
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, J	
Schein, Duane Smailus, Michael Smith, Lorraine Spence, J	
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jil	
George Yang, Scott Zuyderduyn, Marco Marra.	
Clone distribution: MGC clone distribution information ca	
through the I.M.A.G.E. Consortium/LLNL at: http://image.1	
Series: TRAL Plate: 51 Row: c Column: 23	
This clone was selected for full length sequencing becaus	
passed the following selection criteria: Hexamer frequenc	
analysis.	
FEATURES	
source	
1. 353	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="IMAGE:3932215"	
/tissue_type="Skin, melanotic melanoma, high MDR	
/clone_lib="NIH MGC 62"	
/lab_host="DH10B"	
/note="Vector: pDNR-LIB"	
ORIGIN	
Query Match 1.94; Score 26; DB 9; Length 353;	
Best Local Similarity 100.0%; Pred. No. 0.021;	
Matches 26; Conservative 0; Mismatches 0; Indels 0;	

BD1225/5	BD122575	416 bp	DNA	linear	PAT 18
LOCUS					

d encoded human protein.

75.1 GI:23217520

2010789-A/14652.

apiens (human)

ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ies 1 to 416)

is, J.B.D.M., Jobert, S. and Giordano, J.E.

id encoded human protein

JP 2002010789-A 14652 15-JAN-2002;

1 CORP

homo sapiens (human)

IP 2002010789-A/14652

15-JAN-2002

07-AUG-2000 JP 2000280989

05-AUG-1999 US 60/147499

JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI

DANO

12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

1/21, C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC

3/00

IST and encoded human protein

key

source

Location/Qualifiers

1..416 /organism='Homo sapiens (human)'

Location/Qualifiers

1..416

/organism='Homo sapiens'

/mol_type='genomic DNA'

/db_xref='taxon:9606'

1.9%; Score 26; DB 6; Length 416;

larity 100.0%; Pred. NO. 0.022;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TCATGTAGACTTTGAAATTC 26

TCATGTAGACTTTGAAATTC 356

480

sapiens, clone IMAGE:5214272, mRNA, partial cds.

480

480.1 GI:21619102

sapiens (human)

sapiens

yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

lia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ases 1 to 418)

usberg, R.

it Submission

ated (06-JUN-2002) National Institutes of Health, Mammalian

Collection (MGC), Cancer Genomics Office, National Cancer

tute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

4GC Project URL: <http://mgc.nci.nih.gov>

act: MGC help desk

l: cgapbs-remail.nih.gov

ie Procurement: Life Technologies, Inc.

Library Preparation: Life Technologies, Inc.

Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

Sequencing by: National Institutes of Health Intramural

encing Center (NISC),

ersburg, Maryland;

site: <http://www.nisc.nih.gov/>

act: nisc_mgc@nhgri.nih.gov

Akhter, N., Ayele, K., Becketrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Broc,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Lega,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McClos,
McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchi,
Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggir
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information car
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov/>
Series: IRAC plate: 64 Row: 1 Column: 24
This clone was selected for full length sequencing because
passed the following selection criteria: Hexamer frequency
analysis.

FEATURES

Location/Qualifiers

1..418

/organism='Homo sapiens'

/mol_type='RNA'

/db_xref='taxon:9606'

/clone='IMAGE:5214272'

/tissue_type='Blood, adult leukocytes'

/clone_lib='NIH MGC_118'

/lab_host='DH10B'

/note='Vector: pCMV-SPORT6'

41..281

/codon_start=3

/product='Unknown (protein for IMAGE:5214272)'

/protein_id='AAH32480.1'

/db_xref='GI:21619103'

/translation='RSVLLLVAVRLHLLSCPLQBPAGTEWILEEGV
DIYNLRSPPENWRNMGALWKEKDRPCAFMKVKIWLNFHKVTVVIA

Query Match 1.9%; Score 26; DB 9; Length 418;
Best Local Similarity 100.0%; Pred. NO. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

ORIGIN

QY 1 ATGTCATTGTTAGACTTTGAAATTC 26

DB 309 ATGTCATTGTTAGACTTTGAAATTC 334

RESULT 45

AX330518/c

LOCUS

AX330518 439 bp DNA linear PAT 1

Sequence 1027 from Patent WO0194629.

ACCESSION

AX330518

VERSION

AX330518.1 GI:18103496

KEYWORDS

Source

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

REFERENCE

1

Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, C.

Horrigan, S., Soppet, D.R. and Weaver, Z.

Cancer gene determination and therapeutic screening using

gene sets

Patent: WO 0194629-A 1027 13-DEC-2001;

Avalon Pharmaceuticals (US)

Location/Qualifiers

1..439

/organism='Homo sapiens'

/mol_type='unassigned DNA'

/db_xref='taxon:9606'

FEATURES

source

ORIGIN

Query Match 1.9%; Score 26; DB 6; Length 439;

Best Local Similarity 100.0%; Pred. NO. 0.022;

Matches 26; Conservative 0; Mismatches 0; Indels 0;

QY 1 ATGTCATTGTTAGACTTTGAAATTC 26

16:25:14 2004

us-09-245-198a-3.oligo.rge

F

|||||
ATTGTTAGACTTTGAAATTC 47

17 439 bp DNA PAT 14-JUN-2002
e 2744 from Patent WO0229103.
17
17.1 GI:21442802

apiens (human)
ata, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
3, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
pression profiles in liver cancer
e WO 0229103-A 2744 11-APR-2002;
GIC INC (US)
Location/Qualifiers
1. .439
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. N98464"

1.9%; Score 26; DB 6; Length 439;
arity 100.0%; Pred. No. 0.022;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

CATTGTTAGACTTTGAAATTC 26
|||||
CATTGTTAGACTTTGAAATTC 47

34 452 bp DNA PAT 18-DEC-2003
ce 10997 from Patent EP1033401.
34
34.1 GI:40050018

apiens (human)
apiens
ota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.
sed sequence tags and encoded human proteins
e EP 1033401-A 10997 06-SEP-2000;
(FR)

Location/Qualifiers
1. .452
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

1.9%; Score 26; DB 6; Length 452;
arity 100.0%; Pred. No. 0.022; Indels 0; Gaps 0;
Conservative 0; Mismatches 0;

TCATTGTTAGACTTTGAAATTC 26
|||||
TCATTGTTAGACTTTGAAATTC 439

667 452 bp DNA PAT 27-AUG-2002
ace tag and encoded human protein.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD030667
BD030667.1 GI:22572409
JP 2001269182-A/6913.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 452)
Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
Sequence tag and encoded human protein
Patent: JP 2001269182-A 6913 02-OCT-2001;
GENSET

COMMENT

OS Homo sapiens (human)
PN JP 2001269182-A/6913
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEA
PC JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/90, C12
G06F15/40

FEATURES

CC Key Location/Qualifiers.

source

1. .452
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.9%; Score 26; DB 6; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 414 ATGTCATTGTTAGACTTTGAAATTC 439

RESULT 49

AX381620

LOCUS

AX381620 483 bp DNA linear PAT 18

DEFINITION

Sequence 558 from Patent WO0212280.

ACCESSION

AX381620

VERSION

AX381620.1 GI:19576442

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

REFERENCE

AUTHORS

Pyle, R.A., Xu, J. and Secrist, H.

TITLE

Compositions and methods for the therapy and diagnosis of
cancer

JOURNAL

Patent: WO 0212280-A 558 14-FEB-2002;
CORIXA CORPORATION (US)

FEATURES

Location/Qualifiers

source

1. .483
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.9%; Score 26; DB 6; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.022; Indels 0;
Matches 26; Conservative 0; Mismatches 0;

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 397 ATGTCATTGTTAGACTTTGAAATTC 422

iard; DNA; 153 BP.
 (first entry)
 ell derived DNA fragment #168.
 ell; immunosuppressive; immunostimulant; antiinflammatory;
 ene therapy; vaccine; allergen; transplant rejection;
 host disease; malignant disease; ds.
 2000DE-01021834.
 2000DE-01021834.
 THERAPEUTICS GMBH.
 inter H, Reinhartz J;
 320/04.
 cative of T cell activation and functional status, useful
 ; and therapy e.g. of autoimmunity or transplant rejection.
 ; 48; 94pp; German.
 ; represents a novel messenger RNA, (mRNA), (I), for use as
 the activation and functional status of T cells, that have
 reduced expression, and are present at higher or lower
 4, in activated T cells, relative to normal or resting cells,
 ridizes to any of 334 sequences, reproduced, or their
 complements or fragments. The products of the invention have
 sive, immunostimulant, antiinflammatory and cytostatic
 can be used for gene therapy. The polynucleotides of the
 : used: (i) as reagent for detecting activation/functional
 cells, for diagnosis, therapy, modulation or control of the
 ses of (auto)immunity (against microorganisms, vaccines or
 ransplant rejection; immunologically-related inflammation;
 sion; immune deficiency; guest versus host disease, and
 eases of the immune system; (ii) for identifying agents,
 rmaceuticals, that bind to (iii) or derived polypeptides
 to prepare kits for measuring gene expression profiles in
 ne, especially T cells; (iv) to raise antibodies (Ab)
 nest (III); and (v) to prepare binding molecules (IV)
 (II). Ab and (IV) are also useful for detecting and
 e activation and functional status of T cells. AAI68865-
 the activated T-cell derived polynucleotide fragments
 the method of the invention
 BP; 43 A; 23 C; 28 G; 59 T; 0 U; 0 Other;
 .arity 1.9%; Score 26; DB 6; Length 153;
 .arity 100.0%; Pred. No. 0.11;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 TCATTGTTAGACTTTGAAATTC 26
 |||||
 TCATTGTTAGACTTTGAAATTC 55
 iard; cDNA; 281 BP.

DT 07-NOV-2001 (first entry)
 XX Human ovarian PCR-subtracted cDNA library clone #890.
 DE Immunogenic protein; cancer; ovarian tumour; T-cell stimulation;
 XX gene therapy; cytostatic; T-cell expansion; nucleic acid hybridis
 KW primer; probe.
 KW Homo sapiens.
 XX WO200157207-A2.
 PN 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US003733.
 XX 04-FEB-2000; 2000US-0180403P.
 PR 28-MAR-2000; 2000US-0192745P.
 XX (CORI-) CORIXA CORP.
 XX Algate PA, Mannion J;
 PI WPI; 2001-488879/53.
 XX New polynucleotides encoding ovarian tumor proteins, useful for t
 PT ovarian cancer, and as probes, primers, and markers of cancer
 PT progression.
 XX Example 1; Page 253; 378pp; English.
 PS The invention comprises compositions used for the therapy and dia
 XX of ovarian cancer. The compositions comprise one or more ovarian
 CC proteins, their associated polynucleotides, or immunogenic portic
 CC the proteins. The ovarian tumour polynucleotides and polypeptides
 CC useful for stimulating and/or expanding T cells specific for a tu
 CC protein. They are also useful for inhibiting the development of c
 CC a patient with an ovarian tumour DNA or protein by incubating the
 CC cells allowing them to proliferate, and administering to the pati
 CC sequences can be used as markers for cancer, for example, to moni
 CC ovarian cancer progression. Probes and primers are useful in nucl
 CC hybridisation, in detecting the presence of complementary sequen
 CC given sample, for preparing mutant species and for preparing oth
 CC genetic constructions. Sequences AAG23820-AAS25231 and AAS25328-
 CC represent human ovarian tumour protein cDNA clones
 XX Sequence 281 BP; 98 A; 49 C; 59 G; 74 T; 0 U; 1 Other;
 SQ Query Match 1.9%; Score 26; DB 4; Length 281;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0;
 QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
 |||||
 DB 220 ATGTCATTGTTAGACTTTGAAATTC 245
 RESULT 40
 ABQ60530
 ID ABQ60530 standard; cDNA; 386 BP.
 XX AC ABQ60530;
 XX DT 02-AUG-2002 (first entry)
 XX DE Human colon cancer related nucleotide sequence SEQ ID NO:4225.
 XX Human; colon cancer; cancer; tissue profiling; forensic; mapping
 KW genetic analysis; diagnostic; antisense therapy; gene; ss.
 XX OS Homo sapiens.
 XX WO200229086-A2.

001WO-US030732.
 000US-0237271P.
 CORP.
 ste JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
 Lewis ME;
 15/45.
 nucleic acid that is differentially expressed in cancer
 for determining the presence of colon cancer in a cell or
 in antisense therapy.
 796pp; English.
 3060787 represent isolated nucleic acids (I) differentially
 cancer tissues. AB878993 to ABB79004 represent proteins
 ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 use therapy. An antibody immunoreactive with a polypeptide
 is useful for detecting cancer in a patient sample, and
 the presence or absence of a polynucleotide encoded by a
 which hybridises to (I) in a cell. A probe/primer derived
 be used for determining the presence of a nucleic acid which
 (I), and for determining the phenotype of cells in a sample
 a patient. (I) is useful for determining the presence of
 in a cell or tissue type, for determining the presence of
 type of cancer, in antisense therapy, to generate
 a solid surface, to identify a chromosome on which the
 gene resides, and in tissue profiling, forensics, genetic
 ing and diagnostic applications. (I) can be used to raise
 id to screen for peptide analogues and antagonists
 3P; 128 A; 72 C; 85 G; 99 T; 0 U; 2 Other;
 1.9%; Score 26; DB 6; Length 386;
 100.0%; Pred. No. 0.11;
 0; Mismatches 0; Indels 0; Gaps 0;
 1ATTGTTAGACTTTGAAATTC 26
 1ATTGTTAGACTTTGAAATTC 343
 iard; cDNA; 391 BP.
 (first entry)
 tumour associated polynucleotide sequence SEQ ID NO:961.
 a tumour; ovarian cancer; diagnosis; gene therapy;
 vaccine; ss.
 2.
 2001WO-US001575.
 2000US-0176722P.
 A CORP.

PI Algate PA;
 XX
 DR WPI; 2001-425866/45.
 XX
 PT Novel ovarian tumor proteins, and nucleic acids encoding them, usi
 PT treat and diagnose cancers, particularly ovarian cancer.
 XX
 PS Claim 5; Page 239; 338pp; English.
 XX
 CC AAH82377 to AAH83678 represent human ovarian tumour-associated
 CC polynucleotide sequences which encode ovarian tumour proteins. The
 CC ovarian tumour protein and polynucleotide sequences have cytostat
 CC activity, and can be used in gene therapy and vaccine production.
 CC ovarian tumour proteins and polynucleotides can be used to inhibi
 CC development of cancer, particularly ovarian cancer. They can also
 CC to diagnose the onset and progression of cancer
 XX
 SQ Sequence 391 BP; 118 A; 80 C; 88 G; 97 T; 0 U; 8 Other;
 Query Match 1.9%; Score 26; DB 5; Length 391;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; G
 QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
 Db 220 ATGTCATTGTTAGACTTTGAAATTC 245
 RESULT 42
 ABX74646/c
 ID ABX74646 standard; cDNA; 425 BP.
 XX
 AC ABX74646;
 XX
 DT 21-MAR-2003 (first entry)
 XX
 DE Human cDNA sequence #113 up-regulated in CC-RCC patients.
 XX
 KW Human; microarray; solid surface; immobilised probe; CC-RCC;
 KW differential expression profile; aggressive CC-RCC tumour type;
 KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
 KW gene expression profiling; tumour tissue; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200279411-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 29-MAR-2002; 2002WO-US009576.
 XX
 PR 29-MAR-2001; 2001US-0279411P.
 XX
 PA (VAND-) VAN ANDEL INST.
 XX
 PI Haab B, Rhodes D, Teh ET, Takashi M;
 XX
 DR WPI; 2003-040679/03.
 XX
 PT New microarray, comprising a matrix of cDNA probe from a set of p
 PT immobilized to a solid surface in predetermined order, useful in
 PT prognosis of patients with clear cell renal carcinoma.
 XX
 PS Claim 35; SEQ ID NO 223; 179pp; English.
 XX
 CC The present invention relates to a microarray comprising a matrix
 CC least one cDNA probe from a set of probes immobilised to a solid
 CC in a predetermined order, where a row of pixels corresponds to re
 CC of one distinct probe from the set. The probes are complementary
 CC nucleic acid sequences that are expressed differentially in aggre
 CC compared to non-aggressive types of clear cell renal carcinoma (C
 CC and that hybridise to the probes under high stringency conditions
 CC microarray is useful for the prognosis of patients with CC-RCC, w

06:25:14 2004

us-09-245-198a-3.oligo.rng

id non-aggressive CC-RCC tumour types are characterised by expression profiles of genes that hybridise with one or more listed on the microarray. The arrays are useful for gene profiling of tumour and normal tissues. The present sequence human cDNA sequence up-regulated in CC-RCC patients

BP; 118 A; 91 C; 78 G; 138 T; 0 U; 0 Other;

arity 100.0%; Score 26; DB 7; Length 425;

conservative 0; Mismatches 0; Indels 0; Gaps 0;

TCATTGTTAGACTTGAATTC 26

TCATTGTTAGACTTGAATTC 45

idard; DNA; 439 BP.

(first entry)

arcinoma related gene sequence SEQ ID NO:1027.

; colon; breast; ovary; oesophagus; kidney; thyroid;
; prostate; pancreas; carcinoma; antitumour; cancerous;
gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

A2.

2001WO-US010838.

2000US-0209473P.

2000US-0209531P.

2000US-0233133P.

2000US-0233617P.

2000US-0234009P.

2000US-0234034P.

2000US-0234052P.

2000US-0234509P.

2000US-0234567P.

2000US-0234923P.

2000US-0234924P.

2000US-0235077P.

2000US-0235082P.

2000US-0235134P.

2000US-0235280P.

2000US-0235637P.

2000US-0235638P.

2000US-0235711P.

2000US-0235720P.

2000US-0235840P.

2000US-0235863P.

2000US-0236028P.

2000US-0236032P.

2000US-0236033P.

2000US-0236034P.

2000US-0236109P.

2000US-0236111P.

2000US-0236842P.

2000US-0236891P.

2000US-0237172P.

2000US-0237173P.

2000US-0237278P.

2000US-0237294P.

PR 02-OCT-2000; 2000US-0237295P.

PR 02-OCT-2000; 2000US-0237316P.

PR 03-OCT-2000; 2000US-0237425P.

PR 03-OCT-2000; 2000US-0237598P.

PR 03-OCT-2000; 2000US-0237604P.

PR 03-OCT-2000; 2000US-0237608P.

PR 03-OCT-2000; 2000US-0237608P.

PR 01-NOV-2000; 2000US-0244867P.

PR 01-NOV-2000; 2000US-0245084P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrig-

PI Soppet DR, Weaver Z;

XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a

PT agent to be tested for anti-neoplastic activity, and determining

PT in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 1027; 44pp; English.

XX The present invention describes a method (M1) for screening for

CC neoplastic agent. The method involves exposing cells to a chemi-

CC to be tested for anti-neoplastic activity, determining a change

CC expression of at least one gene (I) of a signature gene set, whe-

CC comprises a sequence (S) selected from 8447 sequences (given in

CC to ABL70110), or is at least 95% identical to (S), where a chang-

CC expression is indicative of anti-neoplastic activity. (I) has cy

CC activity and can be used in gene therapy. M1 can be used for scr

CC anti-neoplastic agent, and can be used for producing a product w

CC the data collected with respect to the anti-neoplastic agent as

CC of M1, and the data is sufficient to convey the chemical structu

CC properties of the agent. M1 can be used in the treatment of canc

CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian,

CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear

CC cancer, infiltrating ductal cancer, infiltrating lobular cancer,

CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma an

CC tumour

XX Sequence 439 BP; 118 A; 101 C; 83 G; 137 T; 0 U; 0 Other;

XX Query Match 1.9%; Score 26; DB 6; Length 439;

XX Best Local Similarity 100.0%; Pred. No. 0.11;

XX Matches 26; Conservative 0; Mismatches 0; Indels 0;

QY 1 ATGTCATTGTTAGACTTGAATTC 26

DB 72 ATGTCATTGTTAGACTTGAATTC 47

RESULT 44

ABN96246/c

ID ABN96246 standard; DNA; 439 BP.

XX AC ABN96246;

XX 13-AUG-2002 (first entry)

XX Gene #2744 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

XX metastatic liver tumour; cytostatic; expression profile; disease

XX disease progression; drug toxicity; drug efficacy; drug metaboli

XX Homo sapiens.

XX WO200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US030589.

PF

2000US-0237054P.
LOGIC INC.
ares C, Peres-Da-Silva S, Vockley JG;
119/45.
i detecting the progression of liver cancer, hepatocellular
metastatic liver tumor in a patient, involves detecting the
ession of two or more genes in a liver tissue sample.
ID NO 2744; 298pp; English.
relates to a novel method for diagnosing and detecting the
f liver cancer, hepatocellular carcinoma or metastatic liver
atient, and differentiating metastatic liver cancer from
r carcinoma in a patient, involving detecting the level of
two or more genes represented in ABN93503-ABN97455 in a
The method of the invention has hepatotropic, and
ivity. The method is useful for diagnosing and detecting
on of liver cancer, hepatocellular carcinoma and metastatic
na in a patient. The method is useful for identifying
files which serve as useful diagnostic markers as well as
can be used to monitor disease states, disease progression,
drug efficacy and drug metabolism. Note: The sequence data
at did not form part of the printed specification, but was
electronic format directly from WIPO at
pub/published_pct_sequences
BP; 118 A; 101 C; 83 G; 137 T; 0 U; 0 Other;
1.9%; Score 26; DB 6; Length 439;
arity 100.0%; Pred. No. 0.11;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
CATTTAGACTTTGAAATTC 26
CATTTAGACTTTGAAATTC 47
iard; cDNA; 452 BP.
(first entry)
i protein 5' EST, SEQ ID NO: 10997.
; expressed sequence tag; secreted protein; cDNA isolation;
chromosome mapping; ss.
2000EP-00200610.
99US-0122487P.
F.
iwards J, Duclert A, Giordano J;
381/45.
id that is a 5' expressed sequence tag (5' EST) for
is and genomic DNAs that correspond to 5'ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedu
XX Claim 1; SEQ ID NO 10997; 71pp + Sequence Listing; English.
PS
XX
CC The present sequence is one of a large number of 5' ESTs derived
CC mRNAs encoding secreted proteins. NO ORF has yet been conclusivel
CC identified within the present sequence. The 5' ESTs were prepared
CC total human RNAs or polyA+ RNAs derived from 30 different tissues
CC sequences usually correspond mainly to the 3' untranslated region
CC of the mRNA because they are often obtained from oligo-dT primed
CC libraries. Such ESTs are not well suited for isolating cDNA sequ
CC derived from the 5' ends of mRNAs and even in those cases where l
CC cDNA sequences have been obtained, the full 5' UTR is rarely incl
CC ESTs are derived from mRNAs with intact 5' ends and can therefore
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also us
CC diagnostic, forensic, gene therapy and chromosome mapping procedu
CC They are used to obtain upstream regulatory sequences and to desi
CC expression and secretion vectors
XX
SQ Sequence 452 BP; 122 A; 95 C; 112 G; 122 T; 0 U; 1 Other;
Query Match 1.9%; Score 26; DB 3; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G
QY 1 ATGTCATTGTTAGACTTTGAAATTC 26
Db 414 ATGTCATTGTTAGACTTTGAAATTC 439
RESULT 46
ABV86720
ID ABV86720 standard; cDNA; 469 BP.
XX
AC ABV86720;
XX
DT 13-DEC-2002 (first entry)
XX Human colon cancer related cDNA SEQ ID NO 31.
DE Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine;
XX ss.
KW Homo sapiens.
XX OS
XX WO200258534-A2.
XX PN
XX 01-AUG-2002.
XX PD
XX 16-NOV-2001; 2001WO-US043704.
XX PF
XX 20-NOV-2000; 2000US-0252222P.
PR 06-FEB-2001; 2001US-0267011P.
PR 28-MAR-2001; 2001US-0279670P.
PR 10-JUL-2001; 2001US-0304037P.
XX PR
PA (CORI-) CORIXA CORP.
XX
XX Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;
PI WPI; 2002-608400/65.
XX DR
XX New isolated tumor colon polynucleotide and polypeptide, useful f
PT diagnosis, prevention and/or treatment of cancer, in particular c
PT cancer.
XX
XX Claim 1; SEQ ID NO 31; 266pp + Sequence Listing; English.
PS The invention relates to a human colon tumour expressed polynucle
XX (i) encoding a polypeptide (ii, ABF67991-ABF67996) comprising: (i
CC 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
CC complements of (i); (iii) at least 20 contiguous residues of (i);
CC sequences that hybridize to (i), under moderately stringent condi

s having at least 75% or 90% identity to (i); or (vi) variants of (i). The compositions and methods of the present invention are useful for the diagnosis, prevention and/or treatment of colorectal cancer. (i) can be used in gene therapy and are useful in pharmaceutical compositions such as vaccines. Sequence data for this patent did not form part of the printed matter, but was obtained in electronic format directly from WIPO int/pub/published_pct_sequences

BP; 142 A; 94 C; 111 G; 122 T; 0 U; 0 Other;

1.9%; Score 26; DB 6; Length 469;

larity 100.0%; Pred. No. 0.11; 0; Indels 0; Gaps 0;

Conservative 0; Mismatches 0;

TCATTGTTAGACTTTGAAATTC 26

TCATTGTTAGACTTTGAAATTC 425

ndard; cDNA; 483 BP.

(first entry)

cancer-associated cDNA, SEQ ID No 558.

cancer; immunogenic; vaccine; tumour; gene; ss.

A2.

2001WO-US023826.

2000US-0223265P.

2000US-0237406P.

2001US-0277495P.

2001US-0302702P.

KA CORP.

J, Secrist H;

7462/30.

cleotide encoding colon tumor polypeptides, useful as treating colon cancers.

309; 425pp; English.

a relates to isolated polynucleotides (I) encoding colon epidermis (II). (I) is useful for stimulating an immune patient and treating colon cancer in a patient. (II) is derived from (I) and (II) are useful for determining the presence of a patient. (I) and (II) are useful in pharmaceuticals, e.g. vaccines, and other compositions for the diagnosis and treatment of colon cancer. A composition comprising a first component physiologically acceptable carriers and immunostimulants, en-presenting cell expressing (II) is useful for inhibiting of cancer in a patient. (I) is useful in the design and of ribozyme molecules for inhibiting expression of tumour and (I). ABK54531-ABK5464 represent human colon cancer cDNA the invention

BP; 155 A; 97 C; 112 G; 119 T; 0 U; 0 Other;

1.9%; Score 26; DB 6; Length 483;

Best Local Similarity 100.0%; Pred. No. 0.11; Matches 26; Conservative 0; Mismatches 0; Indels 0;

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 397 ATGTCATTGTTAGACTTTGAAATTC 422

RESULT 48

AAZ51563

ID AAZ51563 standard; cDNA; 486 BP.

XX AAZ51563;

AC AAZ51563;

XX 21-JUN-2000 (first entry)

XX Human hypoxia response regulating gene, 77H4 related cDNA clone

XX Hypoxia response regulating gene; gene 77H4; human; EST 3D; angi

XX cardiant; apoptosis; vasotropic; cytostatic; ophthalmological; s

XX cerebroprotective; antagonist; regulator; inhibitor; treatment;

XX hypoxia associated pathology; HAP; gene therapy; diagnosis; reti

XX steroid receptor coactivator; SRA; ischaemia; myocardial infarct

XX Homo sapiens.

XX Key Location/Qualifiers

XX polyA_signal 449..454

XX /*tag= a

XX WO200012525-A1.

XX 09-MAR-2000.

XX 27-AUG-1999; 99WO-US020394.

XX 27-AUG-1998; 98US-0098158P.

XX 05-MAY-1999; 99US-0132684P.

XX (QUAR-) QUARK BIOTECH INC.

XX (KOHN/) KOHN K.

XX Einat P, Skaliter R, Feinstein E;

XX WPI; 2000-256577/22.

XX Novel polynucleotides capable of regulating angiogenesis or apop

XX useful for diagnosis and treatment of hypoxia, ischemia and tumo

XX Claim 1; Fig 7b; 78pp; English.

XX The present sequence is the human hypoxia response regulating ge

XX related cDNA clone 3D. The gene 77H4 has similarity to steroid r

XX transcriptional co-activator, SRA function and can serve as a co

XX in some transcriptional complexes. It has vasotropic, cardiant,

XX ophthalmological, cytostatic and cerebroprotective activity. Ant

XX of the encoded protein, functions as a regulator of apoptosis or

XX angiogenesis. The protein encoded by this polynucleotide, the

XX biologically active product from enzymatic activity of the prote

XX inhibitor of the enzymatic activity is useful for regulating hyp

XX associated pathologies (HAP). It is useful for gene therapy, dia

XX and treatment of tumour growth and ischaemia, e.g., retinopathy,

XX myocardial infarction and stroke

XX SQ Sequence 486 BP; 160 A; 92 C; 113 G; 121 T; 0 U; 0 Other;

XX Query Match 1.9%; Score 26; DB 3; Length 486;

XX Best Local Similarity 100.0%; Pred. No. 0.11;

XX Matches 26; Conservative 0; Mismatches 0; Indels 0;

Qy 1 ATGTCATTGTTAGACTTTGAAATTC 26

Db 400 ATGTCATTGTTAGACTTTGAAATTC 425

DE	Human endothelial cell cDNA #4820.
XX	
KW	Human; ss; sequencing by hybridisation; SBH; expressed sequence t
KW	genome mapping; biodiversity; genetic disorder.
XX	
OS	Homo sapiens.
XX	
FN	US2003073623-A1.
XX	
PD	17-APR-2003.
XX	
PF	30-JUL-2001; 2001US-00918995.
XX	
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R T.
PA	(LABA/) LABAT I.
PA	(STAC/) STACHE-CRAIN B.
PA	(DICK/) DICKSON M C.
PA	(JONE/) JONES L W.
XX	
PI	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX	
DR	WPI; 2003-615964/58.
XX	
XX	New polynucleotide sequences obtained from various cDNA libraries
PT	as hybridization probes, as oligomers for PCR, for chromosome and
PT	mapping, in the recombinant production of protein, or in generati
PT	antisense DNA or RNA.
XX	
XX	Claim 1; SEQ ID NO 23899; 44pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising an
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequ
CC	determined by the technique of SBH (sequencing by hybridisation).
CC	included is a purified polypeptide comprising a sequence correspo
CC	a reading frame of the novel polynucleotide. The nucleic acid seq
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human
CC	in forensics, in assessing biodiversities, or in identifying muta
CC	responsible for genetic disorders and other traits. The nucleotid
CC	sequences are also useful as hybridisation probes, as oligomers f
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified poly
CC	is useful for generating antibodies specific for it. The present
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequen
CC	for this patent did not form part of the printed specification, b
CC	obtained in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?DocID=20030073623
XX	
SQ	Sequence 498 BP; 152 A; 119 C; 117 G; 107 T; 0 U; 3 Other;
	Query Match 1.9%; Score 26; DB 8; Length 498;
	Best Local Similarity 100.0%; Pred. No. 0.11;
	Matches 26; Conservative 0; Mismatches 0; Indels 0; G
QY	1 ATGTCATTGTTAGACTTTGAATTC 26
DB	349 ATGTCATTGTTAGACTTTGAATTC 374
	RESULT 51
	AAC01272
ID	ID AAC01272 standard; cDNA; 516 BP.
AC	AAC01272;
XX	
XX	06-OCT-2000 (first entry)
XX	
DE	Human secreted protein 5' EST, SEQ ID NO: 1270.
XX	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA iso
KW	gene therapy; chromosome mapping; ss.

[illegible]

```

03-JUL-2001; 2001US-0302702P.
(CORI-) CORIXA CORP.
Pyle RA, Xu J, Secretist H;
WPI; 2002-257462/30.
Novel polynucleotide encoding colon tumor polypeptides, useful as
vaccines for treating colon cancers.
Claim 1; Page 403; 425pp; English.
The invention relates to isolated polynucleotides (I) encoding cc
tumour polypeptides (II). (I) is useful for stimulating an immune
response in a patient and treating colon cancer in a patient. the
Oligonucleotides derived from (I) are useful for determining the
of cancer in a patient. (I) and (II) are useful in pharmaceutical
compositions, e.g. vaccines, and other compositions for the diag
treatment of colon cancer. A composition comprising a first compo
selected from physiologically acceptable carriers and immunostim
and an antigen-presenting cell expressing (II) is useful for inhi
development of cancer in a patient. (I) is useful in the design a
preparation of ribozyme molecules for inhibiting expression of t
polypeptides and (I). ABK54531-ABK55464 represent human colon ca
sequences of the invention
Sequence 531 BP; 153 A; 105 C; 128 G; 143 T; 0 U; 2 Other;
Query Match 1.9%; Score 26; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 26; Conservative 0; Mismatches 0; Indels 0; G
QY 1 ATGTCATTGTTAGACTTTGAAATTTC 26
459 AAGTCATTGTTAGACTTTGAAATTTC 484
DB
RESULT 53
AAC01271
ID AAC01271 standard; CDNA; 540 BP.
AC AAC01271;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 1269.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA is
XX gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-00200610.
XX 26-FEB-1999; 99US-0122487P.
XX (GEST ) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX P-PSDB; AAG01265.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and fo
XX diagnostic, forensic, gene therapy and chromosome mapping procedu
XX Claim 1; SEQ ID NO 1269; 71pp + Sequence Listing; English.

```

sequence is one of a large number of 5' ESTs derived from 3 secreted proteins. An ORF has been identified within the 5' ESTs were prepared from total human RNAs or polyA+ RNAs 30 different tissues. EST sequences usually correspond to 3' untranslated region (UTR) of the mRNA because they are derived from oligo-dT primed cDNA libraries. Such ESTs are not or isolating cDNA sequences derived from the 5' ends of a full 5' UTR is rarely included. 5' ESTs are derived from intact 5' ends and can therefore be used to obtain full length cDNA sequences. 5' ESTs are also used in diagnostic, forensic, and chromosome mapping procedures. They are used to obtain latory sequences and to design expression and secretion

3P; 135 A; 110 C; 137 G; 152 T; 0 U; 6 Other;

arity 100.0%; Score 26; DB 3; Length 540;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

TATGTTAGACTTTGAAATTC 26

TATGTTAGACTTTGAAATTC 527

lard; cDNA; 570 BP.

first entry)

ssociated gene sequence SEQ ID NO:292.

associated gene; cancer antigen; detection; cancer; ostatic; proliferative; vulnery; immunomodulator; antiaesthatic; antirheumatic; antiarthritic; antiviral; ry; antithyroid; antiallergic; antibacterial; cardiant; i; neuroprotective; thrombolytic; coagulant; neutropic; tipisoriatic; antiangiogenic; gene therapy; inflammation; r; haematopoietic cell disorder; autoimmune disorder; ion; graft versus host disease; organ rejection; hrombolytic; cardiovascular disorder; infection; isease; drug screening; ss.

000WO-US005882.

99US-0124270P.

GENOME SCI INC.

en SM;

33/55.

89.

nucleic acids comprising sequences encoding peptides ating or diagnosing e.g. cancer.

853; 2352pp; English.

C78448 encode the human cancer associated proteins given in B44239. The proteins can have activities based on the lls the genes are expressed in. Example of activities

CC include: cytostatic; proliferative; vulnery; immunomodulator; CC antiidiabetic; antiasthmatic; antirheumatic; antiarthritic; CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiv CC dermatological; neuroprotective; cardiant; thrombolytic; coagular CC neutropic; vasotropic; antipsoriatic and antiangiogenic. The CC polynucleotides and polypeptides can be used for preventing, trea CC ameliorating medical conditions and diagnosing pathological condi CC Polynucleotides, polypeptides, antibodies, agonists and antagonis CC the present invention may be used to treat immune disorders by ac CC or inhibiting the proliferation, differentiation or mobilisation CC immune cells, to treat disorders of haematopoietic cells, autoimm CC disorders, allergic reactions, graft versus host disease and orga CC rejection; modulate haemostatic or thrombolytic activity, modulat CC inflammation, cancers, cardiovascular disorders, neurological dis CC bacterial or viral infections. The peptides, nucleotides, antibod CC agonists and antagonists may be also be used in drug screens. AAC CC AAC78457 and AAB44240 represent sequences used in the exemplifica CC the present invention XX

SQ Sequence 570 BP; 171 A; 112 C; 139 G; 145 T; 0 U; 3 Other;

Query Match 1.9%; Score 26; DB 3; Length 570;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 26; Conservative 0; Mismatches 0; Indels 0; G

QY 1 ATGTCATTTGTTAGACTTTGAAATTC 26

Db 469 ATGTCATTTGTTAGACTTTGAAATTC 494

RESULT 55

AAZ51562

ID RAZ51562 standard; cDNA; 580 BP.

XX AAZ51562;

XX 21-JUN-2000 (first entry)

Human hypoxia response regulating gene, 77H4 related cDNA clone 1;

XX Hypoxia response regulating gene; gene 77H4; human; EST 18E; card. KW apoptosis; angiogenesis; vasotropic; cytostatic; ophthalmological KW cerebroprotective; antagonist; regulator; inhibitor; treatment; ti KW hypoxia associated pathology; HAP; gene therapy; diagnosis; ischa KW steroid receptor coactivator; SRA; retinopathy; myocardial infarct KW stroke; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT polyA_signal 536..541

FT /*tag= a

XX WO200012525-A1.

XX 09-MAR-2000.

XX 27-AUG-1999; 99WO-US020394.

XX 27-AUG-1999; 98US-0098158P.

XX 05-MAY-1999; 99US-0132684P.

XX (QUAR-) QUARK BIOTECH INC.

XX (KOHN/) KOHN K.

XX Einat P, Skaliter R, Feinstein E;

XX WPI; 2000-256577/22.

XX Novel polynucleotides capable of regulating angiogenesis or apoptc PT useful for diagnosis and treatment of hypoxia, ischemia and tumor XX Claim 1; Fig 7a; 78pp; English.

06:25:14 2004

us-09-245-198a-3.oligo.rng

nd to screen for peptide analogues and antagonists

BP; 100 A; 80 C; 61 G; 138 T; 0 U; 21 Other;

1.8%; Score 25; DB 6; Length 400;

arity 100.0%; Pred. No. 0.32;

conservative 0; Mismatches 0; Indels 0; Gaps 0;

CAATGTTAGACTTTGAAATTT 25

|||||

CAATGTTAGACTTTGAAATTT 69

dard; cDNA; 626 BP.

(first entry)

novel human diagnostic protein #230.

some mapping; gene mapping; gene therapy; forensic;
nt; medical imaging; diagnostic; genetic disorder; ss.

2.

2001WO-US008631.

2000US-00540217.

2000US-00649167.

INC.

Liu C, Tang YT;

362/73.

339.

polynucleotide and encoded polypeptides, useful in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits and to assess

CD NO 230; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
is useful as hybridisation probes, polymerase chain
primers, oligomers, and for chromosome and gene mapping,
nant production of (II). The polynucleotides are also used
as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal
(I) or to treat disease states involving (II). (II) is
erating antibodies against it, detecting or quantitating a
tissue, as molecular weight markers and as a food
(I) and its binding partners are useful in medical imaging
using (II). (I) and (II) are useful for treating disorders
rant protein expression or biological activity. The
id polynucleotide sequences have applications in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits to assess biodiversity
other types of data and products dependent on DNA and
quences. AAS4197-AAS4564 represent novel human diagnostic
es of the invention. Note: The sequence data for this
appear in the printed specification, but was obtained in
mat directly from WIPO at
ub/published_pct_sequences

SQL Sequence 626 BP; 173 A; 126 C; 157 G; 170 T; 0 U; 0 Other;

Query Match

1.8%; Score 25; DB 5; Length 626;

Best Local Similarity 100.0%; Pred. No. 0.31;

Matches 25; Conservative 0; Mismatches 0; Indels 0; C

QY 1 ATGTCATTTGTAGACTTTGAAATTT 25

|||||

542 ATGTCATTTGTAGACTTTGAAATTT 566

RESULT 59

AAX23451/c

ID AAX23451 standard; DNA; 24 BP.

XX AC AAX23451;

XX 18-JUN-1999 (first entry)

XX Human TNRL3 RACE primer 2.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF;
XX developmental abnormality; gestational abnormality; prostate ca
XX APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
XX cytoplasmic domain; immunogen; antibody preparation; breast carci
XX apoptosis; human; primer; ss.

OS Synthetic.

OS Homo sapiens.

XX MO9911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US018393.

XX 05-SEP-1997; 97US-00924634.

XX (UNIW) UNIV WASHINGTON.

XX Chaudhary PM;

XX WPI; 1999-20519:/17.

XX New Tumor Necrosis Factor family receptor polypeptides and ligand
XX useful for diagnosis and treatment of prostate cancer and develop
XX or gestational abnormalities.

XX Example VII; Page 121; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) fam
XX receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TN
XX their active fragments. APO4 is useful for diagnosing prostate ca
XX determining levels of APO4 in an individual. Prostate cancer can
XX treated using APO4 selective binding agents linked to a therapeutic
XX moiety. APO4 polypeptides are also useful for identifying selecti
XX binding agents, useful in diagnosis/treatment of disease by bindi
XX agents to the polypeptide/active fragment which is extracellular,
XX expressed on the cell surface. The binding is preferably performe
XX vivo. APO4 polypeptides/ active fragments are also useful for scr
XX for agonists and antagonists by binding and observing the changer
XX activity. Effective pharmacological agents useful in diagnosis or
XX treatment of disease are also identified using APO4 polypeptides/
XX fragments and APO4 signal transducer molecules that specifically
XX with a cytoplasmic domain of APO4 and detecting a change in level
XX activity. The method is performed in vivo or in vitro. APO4 is al
XX are all useful as immunogens for preparing antibodies. APO4 is al
XX useful for diagnosis/treatment of developmental or gestational
XX abnormalities. APO8 was transfected to human breast carcinoma cell
XX MCF-7, and induced apoptosis

XX Sequence 24 BP; 7 A; 9 C; 3 G; 5 T; 0 U; 0 Other;

1.7%; Score 24; DB 2; Length 24;
 larity 100.0%; Pred. No. 1.1;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GATGAGGGGAAGGCTGCTAC 633
 |||||
 GATGAGGGGAAGGCTGCTAC 1

ndard; DNA; 24 BP.

(first entry)

RACE primer 1.

sis factor receptor; signal transducer molecule; TNF; APO4;
 1 abnormality; Gestational abnormality; prostate cancer;
 APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 domain; immunogen; antibody preparation; breast carcinoma;
 uman; primer; ss.

98WO-US018393.

97US-00924634.

WASHINGTON.

5191/17.

rosis Factor family receptor polypeptides and ligands -
 diagnosis and treatment of prostate cancer and developmental
 abnormalities.

Page 121; 156pp; English.

on describes isolated Tumor Necrosis Factor (TNF) family
 peptides: APO4, APO6, APO8 and APO9 or their active
 id isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 fragments. APO4 is useful for diagnosing prostate cancer by
 levels of APO4 in an individual. Prostate cancer can also be
 3 APO4 selective binding agents linked to a therapeutic
 polypeptides are also useful for identifying selective
 is, useful in diagnosis/treatment of disease by binding of
 a polypeptide/active fragment which is extracellular, or
 the cell surface. The binding is preferably performed in
 polypeptides/ active fragments are also useful for screening
 and antagonists by binding and observing the change in APO4
 active pharmacological agents useful in diagnosis or
 disease are also identified using APO4 polypeptides/active
 3 APO4 signal transducer molecules that specifically interact
 laemic domain of APO4 and detecting a change in level of APO4
 a method is performed in vivo or in vitro. APO polypeptides
 all as immunogens for preparing antibodies. APO4 is also
 diagnosis/treatment of developmental or gestational
 3. APO8 was transfected to human breast carcinoma cell line
 induced apoptosis

3P; 7 A; 2 C; 13 G; 2 T; 0 U; 0 Other;

Query Match 1.7%; Score 24; DB 2; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 24; Conservative 0; Mismatches 0; Indels 0;

QY 812 CTGCGCCCTTCTCAGCTACTTCG 835
 |||||
 Db 24 CTGCGCCCTTCTCAGCTACTTCG 1

RESULT 61

AA56003
 ID AAX56003 standard; DNA; 40 BP.

XX

AC AAX56003;

XX

DT 15-JUL-1999 (first entry)

XX

DE Human tumour necrosis factor Apo-3 ligand PCR primer SEQ ID NO:5

XX

KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto

KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;

KW PCR primer; ss.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9919490-A1.

XX

PD 22-APR-1999.

XX

PF 09-OCT-1998; 98WO-US021407.

XX

PR 10-OCT-1997; 97US-0062037P.

PR 17-DEC-1997; 97US-0069862P.

XX

PA (GETH) GENENTECH INC.

XX

PI Ashkenazi AJ, Marsters SA, Pitti R;

XX

WPI; 1999-287982/24.

XX

PT New human Apo3- ligand (a tumor necrosis factor) homologue.

XX

PS Example 2; Page 37; 74pp; English.

XX

CC The present invention describes a human tumour necrosis factor (

CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has

CC cytostatic activity. Apo-3 ligand can be used to induce apoptosi

CC mammalian cancer cells, to induce NF-kappaB-dependent transcript

CC to induce JNK/SAPK-dependent responses in mammalian cells. The p

CC sequence represents an Apo-3 ligand PCR primer, which is used in

CC example from the present invention

XX

SQ Sequence 40 BP; 9 A; 13 C; 13 G; 5 T; 0 U; 0 Other;

Query Match 1.7%; Score 24; DB 2; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 24; Conservative 0; Mismatches 0; Indels 0;

QY. 244 CGGCGATCGCTGTCGCCGAGGAG 267
 |||||
 Db 17 CGGCGATCGCTGTCGCCGAGGAG 40

RESULT 62

ABK40355

ID ABK40355 standard; DNA; 23 BP.

XX

AC ABK40355;

XX

DT 15-JUL-2002 (first entry)

XX

DE Probe for gene amplification analysis of human PRO207.

enign tumour; malignant tumour; lymphoid malignancy;
 urological disorder; stromal disorder; blastocoeleic disorder;
 disorder; immune disorder; angiogenic disorder; cytostatic;
 ve; probe; ss.

1.

2000WO-US003565.

99WO-US005028.

99US-0123972P.

99US-0133459P.

99WO-US012252.

99US-0140650P.

99US-0140653P.

99US-0144758P.

99US-0145698P.

99US-0146222P.

99US-0149395P.

99US-0151689P.

99WO-US020111.

99WO-US021090.

99WO-US028313.

99WO-US028301.

99WO-US028634.

2000WO-US000219.

TECH INC.

Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
 Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
 Wood WI;

567/26.

iclic acids encoding PRO polypeptides, useful for treating
 ngnant tumors, leukemias and lymphoid malignancies,
 angiogenic and immunologic disorders.

ige 140; 302pp; English.

vention relates to the isolation of novel human PRO
 'AAU86128-AAU86162) and the polynucleotide sequences
 The PRO polypeptides, agonists, antagonists or anti-PRO
 ; useful for treating benign or malignant tumours (e.g.
 bladder, breast, etc), leukaemias and lymphoid
 other disorders such as neuronal, glial, astrocytal,
 glandular, macrophagal, stromal and blastocoeleic disorders,
 immune and angiogenic disorders. The polynucleotide
 also useful in gene therapy. The present sequence
 robe used in the methods of the present invention

; 1 A; 7 C; 7 G; 8 T; 0 U; 0 Other;

1.7%; Score 23; DB 6; Length 23;

arity 100.0%; Pred.No. 3.1;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;

TGGGCGCTGTCACGTGTT 1003

|||||

TGGGCGCTGTCACGTGTT 23

lard; DNA; 38 BP.

XX 15-JUL-1999 (first entry)
 DT Human tumour necrosis factor Apo-3 ligand PCR primer SEQ ID NO:6.
 XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto;
 DE NF-kappaB-dependent transcription; JNK/SAPK-dependent response; c
 XX PCR primer; ss.
 KW Synthetic.
 OS Homo sapiens.
 OS WO9919490-A1.
 PN 22-APR-1999.
 XX 09-OCT-1998; 98WO-US021407.
 XX 10-OCT-1997; 97US-0062037P.
 PR 17-DEC-1997; 97US-0069862P.
 XX (GETH) GENENTECH INC.
 PA Ashkenazi AJ, Marsters SA, Pitti R;
 PI WPI; 1999-287982/24.
 DR New human Apo3- ligand (a tumor necrosis factor) homologue.
 XX Example 2; Page 37; 74pp; English.
 CC The present invention describes a human tumour necrosis factor (T
 CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcripti
 CC to induce JNK/SAPK-dependent responses in mammalian cells. The pr
 CC sequence represents an Apo-3 ligand PCR primer, which is used in
 CC example from the present invention
 XX SQ Sequence 38 BP; 9 A; 11 C; 12 G; 6 T; 0 U; 0 Other;
 Query Match 1.6%; Score 22; DB 2; Length 38;
 Best Local Similarity 100.0%; Pred.No. 8.9;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; G
 QY 834 CGGACTCTTCCAGGTTCACTGA 855
 Db 38 CGGACTCTTCCAGGTTCACTGA 17
 RESULT 64
 AAL18265
 ID AAL18265 standard; cDNA; 223 BP.
 XX AAL18265;
 AC 07-DEC-2001 (first entry)
 XX DE Human breast cancer expressed polynucleotide 10722.
 XX Human; breast cancer; cell marker; cytostatic; ss.
 OS Homo sapiens.
 XX WO200151628-A2.
 PN 19-JUL-2001.
 PD 10-JAN-2001; 2001WO-US000798.
 PF 14-JAN-2000; 2000US-0176077P.
 PR 14-MAR-2000; 2000US-0189167P.
 PR 24-MAR-2000; 2000US-0192099P.

2000US-0193480P.
2000US-0205230P.
2000US-0211315P.
2000US-0220534P.

ENNUIUM PREDICTIVE MEDICINE INC.

u Y, Wang Y, Steinmann K;
1856/48.

useful as a marker for the diagnosis of breast cancer.

1912-1913; 3695pp; English.

n relates to human breast cancer expressed polynucleotides (ADA07544-ADA26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic

BP; 46 A; 51 C; 48 G; 78 T; 0 U; 0 Other;

larity 100.0%; Score 22; DB 4; Length 223;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TATTATTTTATTATTATT 1351
|||||
TATTATTTTATTATTATT 119

ndard; cDNA; 263 BP.

(first entry)

cancer expressed polynucleotide 1278.

t cancer; cell marker; cytostatic; ss.

A2.

2001WO-US000798.

2000US-0176077P.
2000US-0189167P.
2000US-0192099P.
2000US-0193480P.
2000US-0205230P.
2000US-0211315P.
2000US-0220534P.

ENNUIUM PREDICTIVE MEDICINE INC.

u Y, Wang Y, Steinmann K;
1856/48.

useful as a marker for the diagnosis of breast cancer.

278; 3695pp; English.

XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX

The invention relates to human breast cancer expressed polynucleotides (ADA07544-ADA26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

SQ Sequence 263 BP; 52 A; 63 C; 63 G; 85 T; 0 U; 0 Other;

Query Match 1.6%; Score 22; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0;

QY 1330 AGATATTTTATTATTATT 1351

|||||
137 AGATATTTTATTATTATT 158

Db

RESULT 66

ADA02995

ID ADA02995 standard; cDNA; 1005 BP.

AC ADA02995;

XX 06-NOV-2003 (first entry)

DE Mouse Sept9 carcinoma associated coding sequence, SEQ ID NO:1513
XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug sc
KW gene; ss.

XX Mus sp.

XX WO2003057146-A2.

XX 17-JUL-2003.

XX 26-DEC-2002; 2002WO-US041414.

XX 26-DEC-2001; 2001US-00035832.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-587068/55.

PT New recombinant nucleic acid encoding carcinoma associated prote
PT useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1513; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) n
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them.
XX invention also encompasses expression vectors and host cells com
XX CA nucleic acid, a polypeptide (especially an antibody) that spe
XX binds to the protein, and a biochip comprising CA nucleic acid o
XX fragments thereof. The sequences of the invention were identifie
XX oncogenic retroviruses, which insert into the genome of the host
XX at random. Many of these do not carry transduced host oncogenes
XX pathogenic trans-acting viral genes, meaning that cancer inciden
XX direct consequence of the effects of proviral integration into h
XX protooncogenes. The CA nucleic acid sequences can be used to dia
XX carcinoma (especially breast cancer, prostate cancer, lymphoma o
XX leukaemia) or a propensity to carcinoma by determination of the
XX of a CA gene, or by determination of CA gene expression in part
XX tissues. CA nucleic acids, proteins and antibodies are also usefi

agents and in screening and evaluating drug candidates. The
 invention represents a specifically claimed murine CA nucleic acid
 sequence. Note: The complete sequence data for this
 invention is available in the printed specification, but was obtained
 from WIPO at
 pub/published_pct_sequences.

BP; 285 A; 268 C; 273 G; 179 T; 0 U; 0 Other;

arity 1.6%; Score 22; DB 8; Length 1005;

conservative 0; Mismatches 0; Indels 0; Gaps 0;

CCATTATGAAGTTCATC 448

CCATTATGAAGTTCATC 396

dard; cDNA; 1005 BP.

(first entry)

DNA.

ostatic; gene therapy; vaccine; carcinoma; lymphomas;
 asm; adenocarcinoma; sarcoma; gene.

A2.

2001WO-US051291.

2001US-00798586.

2001US-00004113.

2001US-00052482.

2001US-00997722.

2001US-00034650.

S DISCOVERY.

engelhard EK;

337/23.

it nucleic acid, useful for treating carcinomas, lymphomas,
 lasm, adenocarcinoma, or sarcomas.

ID NO 561; 2304pp; English.

relates to a novel recombinant nucleic acid comprising a
 sequence selected from any of the 660 sequences fully defined
 in the invention. A polynucleotide of the invention has cytostatic
 activity and may have a use in gene therapy, or in a vaccine. The
 nucleic acids and polypeptides are useful for treating
 g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 present sequence represents a mouse cDNA of the invention.

BP; 285 A; 268 C; 273 G; 179 T; 0 U; 0 Other;

arity 1.6%; Score 22; DB 9; Length 1005;

conservative 0; Mismatches 0; Indels 0; Gaps 0;

CCATTATGAAGTTCATC 448

CCATTATGAAGTTCATC 396

RESULT 68

ADC85475

ID ADC85475 standard; DNA; 1005 BP.

XX

AC ADC85475;

XX

DT 01-JAN-2004 (first entry)

XX

DE Mouse Sept19 coding sequence.

XX

KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated;
 KW secreted; transmembrane; intracellular; ds.

XX

OS Mus sp.

XX

PN WO2003045230-A2.

XX

PD 05-JUN-2003.

XX

PF 02-DEC-2002; 2002WO-US038582.

XX

PR 30-NOV-2001; 2001US-00997722.

XX

PA (SAGR-) SAGRES DISCOVERY.

XX

PI Morris DW, Engelhard EK;

XX

DR WPI; 2003-513603/48.

XX

PT New recombinant nucleic acid comprising a nucleotide sequence of
 PT the carcinoma-associated (CA) genes, useful for screening for dru
 PT candidates for diagnosing or treating carcinomas.

XX

PS Claim 1; SEQ ID NO 261; 983pp; English.

XX

CC The invention relates to a recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the fully defined carcin
 CC associated (CA) genes from the 50 tables given in the specificati
 CC CA proteins are secreted, transmembrane or intracellular proteins
 CC recombinant nucleic acids are useful for screening for drug candi
 CC for diagnosing or treating carcinomas. Sequences given in ADC8521
 CC ADC85514 represent CA genes of the invention.

XX

SQ Sequence 1005 BP; 285 A; 268 C; 273 G; 179 T; 0 U; 0 Other;

Query Match 1.6%; Score 22; DB 9; Length 1005;

Best Local Similarity 100.0%; Pred. No. 7.6;

Matches 22; Conservative 0; Mismatches 0; Indels 0; G

QY 427 GCAGCCCATTCAGTTCATC 448

Db 375 GCAGCCCATTCAGTTCATC 396

RESULT 69

AAS84907

ID AAS84907 standard; cDNA; 1778 BP.

XX

AC AAS84907;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #20711.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; s;

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

2001WO-US008631.
 2000US-00540217.
 2000US-00649167.
 Q INC.
 Liu C, Tang YT;
 9362/73.
 0720.
 polynucleotide and encoded polypeptides, useful in
 forensics, gene mapping, identification of mutations
 for genetic disorders or other traits and to assess
 ID NO 20711; 103pp; English.
 n relates to isolated polynucleotide (I) and polypeptide (II)
 I) is useful as hybridisation probes, polymerase chain
 R) primers, oligomers, and for chromosome and gene mapping,
 binant production of (II). The polynucleotides are also used
 as expressed sequence tags for identifying expressed
 s useful in gene therapy techniques to restore normal
 (II) or to treat disease states involving (II). (II) is
 generating antibodies against it, detecting or quantitating a
 in tissue, as molecular weight markers and as a food
 (II) and its binding partners are useful in medical imaging
 errant protein expression or biological activity. The
 and polynucleotide sequences have applications in
 forensics, gene mapping, identification of mutations
 for genetic disorders or other traits to assess biodiversity
 ce other types of data and products dependent on DNA and
 sequences. AAS64197-AAS94564 represent novel human diagnostic
 ces of the invention. Note: The sequence data for this
 at appear in the printed specification, but was obtained in
 ormat directly from WIPO at
 /pub/published_pct_sequences
 3 BP; 472 A; 408 C; 371 G; 527 T; 0 U; 0 Other;
 larity 100.0%; Pred. No. 7.3;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 3GATGGGGGGGGCGGTGAGG 84
 |||||
 3GATGGGGGGGGCGGTGAGG 205
 idard; cDNA; 2942 BP.
 (first entry)
 carcinoma associated cDNA, SEQ ID NO:1512.
 ; carcinoma associated; oncogene; carcinoma; cancer; breast;
 mphoma; leukaemia; cytostatic; gene therapy; drug screening;
 A2.
 26-DEC-2002; 2002WO-US041414.
 26-DEC-2001; 2001US-00035832.
 (SAGR-) SAGRES DISCOVERY.
 Morris DW;
 WPI; 2003-587068/55.
 New recombinant nucleic acid encoding carcinoma associated prote
 useful for preparing compositions for treating carcinomas.
 Claim 1; SEQ ID NO 1512; 245pp; English.
 The invention relates to recombinant carcinoma associated (CA) n
 acid sequences from mouse and human (ADA01482-ADA03094), and to
 recombinant carcinoma associated proteins (CAP) encoded by them.
 CC invention also encompasses expression vectors and host cells com
 CA nucleic acid, a polypeptide (especially an antibody) that spe
 CC binds to the protein, and a biochip comprising CA nucleic acid o
 CC fragments thereof. The sequences of the invention were identifie
 CC oncogenic retroviruses, which insert into the genome of the host
 CC at random. Many of these do not carry transduced host oncogenes
 CC pathogenic trans-acting viral genes, meaning that cancer inciden
 CC direct consequence of the effects of proviral integration into h
 CC protooncogenes. The CA nucleic acid sequences can be used to dia
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma o
 CC leukaemia) or a propensity to carcinoma by determination of the
 CC of a CA gene, or by determination of CA gene expression in parti
 CC tissues. CA nucleic acids, proteins and antibodies are also usef
 CC therapeutic agents and in screening and evaluating drug candidat
 CC present sequence represents a specifically claimed murine CA nuc
 CC sequence of the invention. Note: The complete sequence data for
 CC patent did not form part of the printed specification, but was o
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 SQ Sequence 2942 BP; 702 A; 837 C; 780 G; 623 T; 0 U; 0 Other;
 Query Match 1.6%; Score 22; DB 8; Length 2942;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0;
 QY 427 GCAGCCCATTTAGAGTTCATC 448
 Db 552 GCAGCCCATTTAGAGTTCATC 573
 RESULT 71
 ADB72732
 ID ADB72732 standard; mRNA; 2942 BP.
 XX
 AC ADB72732;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Mouse Sept9 mRNA.
 XX
 KW mouse; ss; cytostatic; gene therapy; vaccine; carcinoma; lymphoma
 KW cancer; neoplasm; adenocarcinoma; sarcoma.
 XX
 OS Mus sp.
 XX
 PN WO2003008583-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 26-DEC-2001; 2001WO-US051291.
 XX
 PR 02-MAR-2001; 2001US-00798586.
 PR 23-OCT-2001; 2001US-00004113.

2001US-00052482.
2001US-00997722.
2001US-00034650.

S DISCOVERY.

ngelhard EK;

337/23.

nt nucleic acid, useful for treating carcinomas, lymphomas,
lasm, adenocarcinoma, or sarcomas.

ID NO 560; 2304pp; English.

relates to a novel recombinant nucleic acid comprising a
quence selected from any of the 660 sequences fully defined
ication. A polynucleotide of the invention has cytostatic
may have a use in gene therapy, or in a vaccine. The
nucleic acids and polypeptides are useful for treating
.9. lymphomas, cancers, neoplasm, adenocarcinoma, and
present sequence represents a mouse mRNA of the invention.

BP; 702 A; 837 C; 780 G; 623 T; 0 U; 0 Other;

arity 1.6%; Score 22; DB 9; Length 2942;

onservative 0; Mismatches 0; Indels 0; Gaps 0;

CCCATTTATGAAGTTCATC 448

CCCATTTATGAAGTTCATC 573

iard; DNA; 2942 BP.

(first entry)

RNA sequence.

ne therapy; vaccine; cancer; carcinoma-associated gene; CA;
membrane; intracellular; ds.

12.

2002WO-US038582.

2001US-00997722.

S DISCOVERY.

ngelhard EK;

103/48.

it nucleic acid comprising a nucleotide sequence of any of
associated (CA) genes, useful for screening for drug
diagnosing or treating carcinomas.

ID NO 260; 983pp; English.

relates to a recombinant nucleic acid comprising a
quence selected from any of the fully defined carcinoma-
genes from the 50 tables given in the specification. The
e secreted, transmembrane or intracellular proteins. The

CC recombinant nucleic acids are useful for screening for drug candi
CC for diagnosing or treating carcinomas. Sequences given in ADC8521
CC ADC85514 represent CA genes of the invention.

XX SQ Sequence 2942 BP; 702 A; 837 C; 780 G; 623 T; 0 U; 0 Other;

Query Match

Best Local Similarity 1.6%; Score 22; DB 9; Length 2942;

Matches 22; Conservative 0; Mismatches 0; Indels 0; G

Oy 427 GCAGCCCATTTATGAAGTTCATC 448

Db 552 GCAGCCCATTTATGAAGTTCATC 573

RESULT 73

ADA02993

ID ADA02993 standard; DNA; 50295 BP.

XX AC ADA02993;

XX DT 06-NOV-2003 (first entry)

XX Mouse Sept9 carcinoma associated gene, SEQ ID NO:1511.

XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer;

XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug scr

XX gene; ds.

XX OS Mus sp.

XX WO2003057146-A2.

XX PD 17-JUL-2003.

XX PF 26-DEC-2002; 2002WO-US041414.

XX PR 26-DEC-2001; 2001US-00035832.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX WPI; 2003-587068/55.

XX PT New recombinant nucleic acid encoding carcinoma associated protei
XX useful for preparing compositions for treating carcinomas.

XX PS Claim 1; SEQ ID NO 1511; 245pp; English.

XX CC The invention relates to recombinant carcinoma associated (CA) nu
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them.
XX invention also encompasses expression vectors and host cells comp.
XX CA nucleic acid, a polypeptide (especially an antibody) that spec
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified
XX oncogenic retroviruses, which insert into the genome of the host
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence
XX direct consequence of the effects of proviral integration into hos
XX protooncogenes. The CA nucleic acid sequences can be used to diagn
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukaemia) or a propensity to carcinoma by determination of the se
XX of a CA gene, or by determination of CA gene expression in partic
XX tissues. CA nucleic acids, proteins and antibodies are also useful
XX therapeutic agents and in screening and evaluating drug candidates
XX present sequence represents a specifically claimed murine CA nucle
XX sequence of the invention. Note: The complete sequence data for th
XX patent did not form part of the printed specification, but was obt
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

06:25:14 2004

us-09-245-198a-3.oligo.rng

95 BP; 10891 A; 12938 C; 13611 G; 12237 T; 0 U; 618 Other;
1.6%; Score 22; DB 8; Length 50295;
arity 100.0%; Pred. No. 6.2;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GCCATTATGAAGTTCATC 448
|||||
GCCATTATGAAGTTCATC 31093

ndard; DNA; 50295 BP.

(first entry)
gene.
ytostatic; gene therapy; vaccine; carcinoma; lymphomas;
laem; adenocarcinoma; sarcoma; gene.

-A2.

2001WO-US051291.
2001US-00798586.
2001US-00004113.
2001US-00052482.
2001US-00997722.
2001US-00034650.

ES DISCOVERY.
Engelhard EK;
3337/23.

ant nucleic acid, useful for treating carcinomas, lymphomas,
plasm, adenocarcinoma, or sarcomas.

ID NO 559; 2304pp; English.

relates to a novel recombinant nucleic acid comprising a
sequence selected from any of the 660 sequences fully defined
ication. A polynucleotide of the invention has cytostatic
i may have a use in gene therapy, or in a vaccine. The
nucleic acids and polypeptides are useful for treating
g. lymphomas, cancers, neoplasm, adenocarcinoma, and
; present sequence represents a mouse gene of the invention.

95 BP; 10891 A; 12938 C; 13611 G; 12237 T; 0 U; 618 Other;
1.6%; Score 22; DB 9; Length 50295;
arity 100.0%; Pred. No. 6.2;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GCCATTATGAAGTTCATC 448
|||||
GCCATTATGAAGTTCATC 31093

ndard; DNA; 50295 BP.

XX 01-JAN-2004 (first entry)
DT Mouse Sept19 genomic sequence.
DE
XX Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated
KW secreted; transmembrane; intracellular; ds.
XX Mus sp.
XX WO2003045230-A2.
XX 05-JUN-2003.
XX 02-DEC-2002; 2002WO-US038582.
XX 30-NOV-2001; 2001US-00997722.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW, Engelhard EK;
PI WPI; 2003-513603/48.
DR
XX New recombinant nucleic acid comprising a nucleotide sequence of
PT the carcinoma-associated (CA) genes, useful for screening for dr
PT candidates for diagnosing or treating carcinomas.
XX Claim 1; SEQ ID NO 259; 983pp; English.
XX The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carci
CC associated (CA) genes from the 50 tables given in the specificat
CC CA proteins are secreted, transmembrane or intracellular protein
CC recombinant nucleic acids are useful for screening for drug cand
CC for diagnosing or treating carcinomas. Sequences given in ADC852
CC ADC8514 represent CA genes of the invention.
XX
SQ Sequence 50295 BP; 10891 A; 12938 C; 13611 G; 12237 T; 0 U; 618
Query Match 1.6%; Score 22; DB 9; Length 50295;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0;
QY 427 GCAGCCCATTTATGAAGTTCATC 448
|||||
DB 31072 GCAGCCCATTTATGAAGTTCATC 31093

Search completed: April 8, 2004, 21:02:41
Job time : 662 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 CompuGen Ltd.

in search, using sw model

ril 7, 2004, 17:54:48 ; Search time 20 seconds
(without alignments)
1365.920 Million cell updates/sec

-09-245-198A-4

MSLLDFEISARRLLPLPSRLG.....PWAHLKAAPFLTYRFLQVH 284

pop 60.0 , Gapext 60.0

3366 seqs, 96191526 residues

ts satisfying chosen parameters: 283366

3th: 0

3th: 20000000000

isting first 100 summaries

IR 78:*

pir1:*

pir2:*

pir3:*

pir4:*

the number of results predicted by chance to have a
score than or equal to the score of the result being printed,
as by analysis of the total score distribution.

SUMMARIES

seq	length	DB	ID	Description
1.5	111	2	A85866	hypothetical prote
1.2	733	2	S78376	photosystem I P700
1.8	58	2	A58208	protamine I, 1 - pa
1.8	143	2	G84168	hypothetical prote
1.8	197	2	E72374	hypothetical prote
1.8	220	2	AG3547	bicyclomycin resis
1.8	278	2	D83080	hypothetical prote
1.8	339	2	C71132	hypothetical prote
1.8	372	2	H70813	hypothetical prote
1.8	379	2	E64300	probable cysteine
1.8	381	2	AH3041	formate dehydrogen
1.8	387	2	D84885	conserved hypothet
1.8	397	2	D98244	hypothetical prote
1.8	422	1	A60503	hypothetical prote
1.8	443	2	T17220	sperm-binding glyc
1.8	455	2	AC0347	hypothetical prote
1.8	471	2	A75267	probable membrane
1.8	576	2	E64186	probable transport
1.8	586	2	A41125	probable ATP-bind
1.5	45	2	D58208	gamma-glutamyltran
1.5	50	2	S22582	protamine II-3 - p
1.5	58	2	S34045	protamine 1 - Sagu
1.5	86	2	F87604	protamine - North
1.5	102	2	F87993	hypothetical prote
1.5	115	2	PH1560	protein ZC334.3 [i
1.5	115	2	H83201	Ig heavy chain V r
1.5	118	1	IEEC5B	conserved hypothet
1.5	118	2	AE1753	hypothetical prote
1.5	123	2	AH2707	Orf51 [bacterioph
				conserved hypothet

30	7	2.5	125	2	T27519	hypothetical
31	7	2.5	131	2	I52290	interleukin
32	7	2.5	131	2	E30552	T-cell act;
33	7	2.5	146	2	T37116	probable t;
34	7	2.5	147	2	A71217	hypothetical
35	7	2.5	150	2	T08734	hypothetical
36	7	2.5	157	2	S31078	seed aller
37	7	2.5	157	2	T02664	allergen -
38	7	2.5	157	2	A75567	conserved i
39	7	2.5	157	2	E75530	hypothetical
40	7	2.5	160	2	SS925	allergen R
41	7	2.5	161	1	DNEC17	outer membr
42	7	2.5	161	2	D90651	histone-lik
43	7	2.5	161	2	D85502	hypothetical
44	7	2.5	162	2	T24937	hypothetical
45	7	2.5	162	2	T31173	hypothetical
46	7	2.5	164	2	S76920	hypothetical
47	7	2.5	170	2	S44789	D2007.4 prc
48	7	2.5	174	2	D87638	transcripti
49	7	2.5	180	1	LGST	beta-lactog
50	7	2.5	180	1	LGSH	beta-lactog
51	7	2.5	181	2	B60738	insulin-lik
52	7	2.5	187	2	G85343	phospholipa
53	7	2.5	206	2	S72567	hypothetical
54	7	2.5	230	2	AH0692	conserved h
55	7	2.5	231	2	B64920	probable me
56	7	2.5	231	2	E90921	hypothetical
57	7	2.5	231	2	A85770	hypothetical
58	7	2.5	231	2	C86665	amino acid
59	7	2.5	233	2	S60767	ribonucleas
60	7	2.5	233	2	AB0273	probable me
61	7	2.5	234	2	G85098	H+-transpor
62	7	2.5	235	2	I64174	probable so
63	7	2.5	236	2	A86387	probable cy
64	7	2.5	238	2	H70866	hypothetical
65	7	2.5	240	2	G83208	conserved h
66	7	2.5	243	2	C64124	molybdopter
67	7	2.5	243	2	T25942	hypothetical
68	7	2.5	244	2	A46066	lymphotoxin
69	7	2.5	249	2	B32352	molybdopter
70	7	2.5	249	2	H90741	molybdopter
71	7	2.5	249	2	C85592	molybdopter
72	7	2.5	249	2	AI0602	molybdopter
73	7	2.5	252	2	AI3154	IS21 family
74	7	2.5	252	2	H98132	1stB protei
75	7	2.5	254	2	F82439	molybdopter
76	7	2.5	255	2	AB0182	molybdopter
77	7	2.5	258	2	AI0566	hydroxypyru
78	7	2.5	258	2	G83101	probable ac
79	7	2.5	260	2	AG0007	lipopolysac
80	7	2.5	262	2	F71801	flagellar b
81	7	2.5	262	2	A64718	flagellar b
82	7	2.5	277	2	G75518	probable bel
83	7	2.5	293	2	A97396	hypothetical
84	7	2.5	300	2	AB2614	hypothetical
85	7	2.5	303	2	E87280	hard protei
86	7	2.5	307	2	S23780	nucleic aci
87	7	2.5	307	2	A83466	maltose tra
88	7	2.5	308	2	D95932	probable su
89	7	2.5	310	2	D70328	histidine k
90	7	2.5	312	2	A31846	130K paracr
91	7	2.5	313	2	C75208	sugar trans
92	7	2.5	316	1	A46489	pan-epitheli
93	7	2.5	316	2	C82561	drug tolera
94	7	2.5	318	2	AD3295	sodium/bile
95	7	2.5	319	2	C95927	probable su
96	7	2.5	321	2	B72604	hypothetical
97	7	2.5	325	2	C36718	pyruvate del
98	7	2.5	329	2	T18619	hypothetical
99	7	2.5	333	2	B90172	conserved h
100	7	2.5	343	2	A43577	regulatory f

ALIGNMENTS

ein 23516 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 ichia coli
 101 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 66
 unkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 ck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 33, 2001
 sequence of enterohemorrhagic Escherichia coli O157:H7.
 r: A85480; MUID:21074935; PMID:11206551
 66

nary
 DNA
 <STO>
 s: GB:AE005174; NID:g12516604; PIDN:AAG57389.1; GSPDB:GN00145; UWGP:Z35
 urce: strain O157:H7, substrain EDL933

3.5%; Score 10; DB 2; Length 111;
 larity 100.0%; Pred.No. 0.074;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALACGL 72
 |||||
 ALACGL 49

0 apoprotein A2 - Odontella sinensis chloroplast
 plast Odontella sinensis
 98 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000
 76
 Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
 Rep. 13, 336-342, 1995
 chloroplast genome of a chlorophyll a+c- containing Alga, Odontella sinensis
 c: S78238
 76

nary; nucleic acid sequence not shown; translation not shown

DNA
 <KOW>
 i: EMBL:267753; NID:g1185127; PIDN:CAA91749.1; PID:g1185266
 stide sequence was submitted to the EMBL Data Library, November 1995

ast
 rosystem I p700 apoprotein
 plast; electron transfer; membrane protein; membrane-associated complex

3.2%; Score 9; DB 2; Length 733;
 arity 100.0%; Pred.No. 3.3;
 onservative 0; Mismatches 0; Indels 0; Gaps 0;

LACLG 71
 |||||
 LACLG 341

inted turtle
 ys picta (painted turtle)
 6 #sequence_revision 08-Nov-1996 #text_change 07-May-1999
 8
 nsky, H.E.; Elsev, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.
 , 23547-23557, 1996
 s of reptiles.
 : A58208; MUID:96394458; PMID:8798564

A;Accession: A58208
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-58 <HUN>
 C;Superfamily: sperm histone

Query Match 2.8%; Score 8; DB 2; Length 58;
 Best Local Similarity 100.0%; Pred.No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 42 ORRRGRGG 49
 |||||
 Db 35 ORRRGRGG 42

RESULT 4

G84168
 hypothetical protein Vng0080h [imported] - Halobacterium sp. NRC-1
 C;Species: Halobacterium sp. NRC-1
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: G84168

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shu
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Ma
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt,
 A;Title: Genome sequence of Halobacterium species NRC-1.
 A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: G84168
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-143 <STO>
 A;Cross-references: GB:AE004437; NID:g10579733; PIDN:AAG18715.1; GSPD
 C;Genetics:
 A;Gene: VNG0080H

Query Match 2.8%; Score 8; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred.No. 7.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 61 LGLGLALA 68
 |||||
 Db 55 LGLGLALA 62

RESULT 5

E72374
 hypothetical protein - Thermotoga maritima (strain MSB8)
 C;Species: Thermotoga maritima
 C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jun-1999
 C;Accession: E72374
 R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.;
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.;
 C.M.

Nature 399, 323-329, 1999
 A;Title: Evidence for lateral gene transfer between Archaea and Bacter
 A;Reference number: A72200; MUID:99287316; PMID:10360571
 A;Accession: E72374
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-197 <ARN>

A;Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AAD35
 A;Experimental source: strain MSB8
 C;Genetics:
 A;Gene: TM0469

Query Match 2.8%; Score 8; DB 2; Length 197;
 Best Local Similarity 100.0%; Pred.No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 2 SLDDFEIS 9
 |||||
 Db 135 SLDDFEIS 142

tance protein [imported] - Brucella melitensis (strain 16M)
 a melitensis
 2 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 7
 ; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 sman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 sci. U.S.A. 99, 443-448, 2002
 a sequence of the facultative intracellular pathogen Brucella melitens
 ; AD3252; PMID:11756688
 7
 iry
 A
 <KUR>
 : GB:AE008918; PIDN:AAL53546.1; PID:gl7984455; GSPDB:GN00191
 ce: strain 16M

 2.8%; Score 8; DB 2; Length 220;
 irity 100.0%; Pred. No. 11;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 LAL 61
 ||||
 LAL 139

 n PA4521 [imported] - Pseudomonas aeruginosa (strain PA01)
 nas aeruginosa
 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 m, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
 ; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 M.V.
 , 2000
 enome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A82950; PMID:20437337; PMID:10984043
 iry
 A
 STO>
 GB:AE004866; GB:AE004091; NID:g9950760; PIDN:AAG07909.1; GSPDB:GN001
 ce: strain PA01

 2.8%; Score 8; DB 2; Length 278;
 rity 100.0%; Pred. No. 13;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 AW 77
 ||||
 AW 54

 n PH0824 - Pyrococcus horikoshii
 us horikoshii
 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
 Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 nahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi
 1998
 equence and gene organization of the genome of a hyper-thermophilic a
 A71000; PMID:98344137; PMID:9679194

A;Accession: C71132
 A;Status: preliminary; nucleic acid sequence not shown; translation nc
 A;Molecule type: DNA
 A;Residues: 1-339 <RAW>
 A;Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29917.1; PID:g3
 A;Experimental source: strain OT3
 A;Note: this accession replaces an interim accession for a sequence re
 C;Genetics:
 A;Gene: PH0824
 C;Superfamily: conserved hypothetical protein MTH900

 Query Match 2.8%; Score 8; DB 2; Length 339;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G

 QY 70 LGLLLAVV 77
 |||||
 Db 111 LGLLLAVV 118

 RESULT 9
 H70813
 probable cysteine synthase - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun
 C;Accession: H70813
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Har
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamli
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squ
 Nature 393, 537-544, 1998
 A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barre
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from th
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: H70813
 A;Status: preliminary; nucleic acid sequence not shown; translation no
 A;Molecule type: DNA
 A;Residues: 1-372 <COL>
 A;Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CAA17
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: cysM3
 C;Superfamily: threonine dehydratase

 Query Match 2.8%; Score 8; DB 2; Length 372;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G

 QY 61 LGLGLALA 68
 |||||
 Db 96 LGLGLALA 103

 RESULT 10
 E64300
 formate dehydrogenase (EC 1.2.1.2) beta chain - Methanococcus jannaschi
 C;Species: Methanococcus jannaschi
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-
 C;Accession: E64300
 R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutt
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick,
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurs
 Science 273, 1058-1073, 1996
 A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smit
 A;Title: Complete genome sequence of the methanogenic archaeon, Methano
 A;Reference number: A64300; MUID:96337999; PMID:8688087
 A;Accession: E64300
 A;Status: preliminary; nucleic acid sequence not shown; translation not
 A;Molecule type: DNA
 A;Residues: 1-379 <BU>
 A;Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AAB97986.1
 C;Genetics:
 A;Map position: REV7250-6111
 C;Superfamily: formate dehydrogenase chain B; ferredoxin 2 [4Fe-4S] homo

06:25:20 2004

us-09-245-198a-4.oligo.rpr

```
reductase
ferredoxin 2[4Fe-4S] homology <FER>
    2.8%; Score 8; DB 2; Length 379;
    arity 100.0%; Pred. No. 17;
    Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    3VAL 224
    11111
    3VAL 42
    11111

ical protein Atu3948 [imported] - Agrobacterium tumefaciens (strain C58)
22 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
11
    bal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
    W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
    P.; Zhang, S.
    2323, 2001
    ; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
    ne of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
    C: AB2577; MUID:21608550; PMID:11743193
    11
    ary
    DNA
    <KUR>
    C: GB:AE008689; PIDN:AAL44750.1; PID:G17742385; GSPDB:GN00187
    ource: strain C58 (Dupont)

near chromosome
    2.8%; Score 8; DB 2; Length 381;
    arity 100.0%; Pred. No. 17;
    Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    PGSS 258
    11111
    PGSS 332
    11111

in At2g45000 [imported] - Arabidopsis thaliana
psis thaliana (mouse-ear cross)
11 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
15
    ; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
    t, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
    W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
    8, 1999
    and analysis of chromosome 2 of the plant Arabidopsis thaliana.
    : AB4420; MUID:20083487; PMID:10617197
    5
    ary
    NA
    <STO>
    : GB:AE002093; NID:G4895250; PIDN:AAD32835.1; GSPDB:GN00139

    2.8%; Score 8; DB 2; Length 387;
    arity 100.0%; Pred. No. 18;
    Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    AEED 100
    11111
    AEED 377
```

```
RESULT 13
D98244
hypothetical protein AGR_L1808 [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2001
C:Accession: D98244
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qu
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappa
Science 294, 2323-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agen
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D98244
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89478.1; PID:G15159347; GSPD
C:Genetics:
A:Gene: AGR_L1808
A:Map position: linear chromosome
    Query Match 2.8%; Score 8; DB 2; Length 397;
    Best Local Similarity 100.0%; Pred. No. 18;
    Matches 8; Conservative 0; Mismatches 0; Indels 0;
    QY 251 LALRPGSS 258
    1111111111
    Db 341 LALRPGSS 348

RESULT 14
A60503
sperm-binding glycoprotein ZP3 precursor - Golden hamster
N:Alternate names: sperm receptor; zona pellucida glycoprotein ZP3
C:Species: Mesocricetus auratus (golden hamster)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A60503
R:Kinloch, R.A.; Ruiz-Seiler, B.; Wasserman, P.M.
Dev. Biol. 142, 414-421, 1990
A>Title: Genomic organization and polypeptide primary structure of zo
A:Reference number: A60503; MUID:91078540; PMID:2257975
A:Accession: A60503
A:Molecule type: DNA
A:Residues: 1-422 <KIN>
A:Cross-references: GB:M63629
A>Note: the authors translated the codon CAA for residue 251 as Glu, &
C:Comment: This sulfated glycoprotein in the zona pellucida of the ooc
C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology
C:Keywords: glycoprotein; oocyte
F:45-300/Domain: ZP domain homology <ZPH>
    Query Match 2.8%; Score 8; DB 1; Length 422;
    Best Local Similarity 100.0%; Pred. No. 19;
    Matches 8; Conservative 0; Mismatches 0; Indels 0;
    QY 59 LALGLGLA 66
    1111111111
    Db 386 LALGLGLA 393

RESULT 15
T17220
hypothetical protein DKFZp5660011.1 - human
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun
C:Accession: T17220
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18725
A:Accession: T17220
A>Status: preliminary
A:Molecule type: mRNA
```

<BLU>
: EMBL:AL117414
rce: fetal kidney; clone DKFZp5660011

11.1
na-glutamyltransferase

2.8%; Score 8; DB 2; Length 443;
arity 100.0%; Pred. No. 20;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

LALA 68
|||||
LALA 21

protein yegB [imported] - Yersinia pestis (strain CO92)
a pestis
1 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
en, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
7, 2001
ence of Yersinia pestis, the causative agent of plague.
: AB0001; MUID:21470413; PMID:11586360

ry
JA
:KUR>

: GB:AL590842; PIDN:CAC92102.1; PID:g15980820; GSPDB:GN00175

idrug-efflux transporter

2.8%; Score 8; DB 2; Length 465;
arity 100.0%; Pred. No. 21;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

VSL 79
|||||
VSL 342

protein - Deinococcus radiodurans (strain R1)

#sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
athevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
nter, J.C.; Fraser, C.M.
577, 1999

ence of the radioresistant bacterium Deinococcus radiodurans R1.
A75250; MUID:20036896; PMID:10567266

ry
A
WHI>

GB:AE002079; GB:AE000513; NID:g6460315; PIDN:AAF12043.1; PID:g646032
ce: strain R1

2.8%; Score 8; DB 2; Length 471;
arity 100.0%; Pred. No. 21;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

LAL 67

Db 366 ALGLGLAL 373
|||||

RESULT 18
E64186

probable ATP-binding transport protein H1156 - Haemophilus influenzae
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb
C:Accession: E64186

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness,
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Ge
Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smi
A:Title: Whole-genome random sequencing and assembly of Haemophilus in
A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: E64186
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-576 <TIGR>

A:Cross-references: GB:U32795; GB:I42023; NID:g1574708; PIDN:AAC22811.
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding c

C:Keywords: ATP; nucleotide binding; P-loop

F:355-550/Domain: ATP-binding cassette homology <ABC>

F:372-379/Region: nucleotide-binding motif A (P-loop)

Query Match 2.8%; Score 8; DB 2; Length 576;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 58 PLALGLGL 65
|||||

Db 159 PLALGLGL 166
|||||

RESULT 19

A41125

gamma-glutamyltransferase (EC 2.3.2.2) related protein - human

N;Alternate names: gamma-glutamyltransferase-like activity 1; GGT-REL

C:Species: Homo sapiens (man)

C:Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 18-Jun

C:Accession: A41125

R;Heisterkamp, N.; Rajpert-De Meyts, E.; Uribe, L.; Forman, H.J.; Grof.
Proc. Natl. Acad. Sci. U.S.A. 88, 6303-6307, 1991

A:Title: Identification of a human gamma-glutamyl cleaving enzyme rela

A:Reference number: A41125; MUID:91296809; PMID:1676842

A:Accession: A41125

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-586 <HEI>

A:Cross-references: GB:M64099; NID:g183141; PIDN:AA58503.1; PID:g1831.

C:Genetics:

A:Gene: GDB:GGTAL1; GGT-REL

A:Cross-references: GDB:134033

C:Superfamily: gamma-glutamyltransferase

C:Keywords: aminocyltransferase; glycoprotein; transmembrane protein

Query Match 2.8%; Score 8; DB 2; Length 586;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Ga

QY 61 LGLGLALA 68
|||||

Db 14 LGLGLALA 21
|||||

RESULT 20

D58208

protamine II-3 - painted turtle

C:Species: Chrysemys picta (painted turtle)

C:Date: 08-Nov-1996 #sequence_revision 08-Nov-1996 #text_change 07-May-
C:Accession: D58208

06:25:20 2004

us-09-245-198a-4.oligo.rpr

insky, H.E.; Elsev, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.
1, 23547-23557, 1996
ies of reptiles.
r: A58208; MUID:96394458; PMID:8798564
08
nary
protein
<HUN>
erm histone
2.5%; Score 7; DB 2; Length 45;
larity 100.0%; Pred. No. 26;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GRRG 49
|||||
GRRG 30
uinus imperator
us imperator
93 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
82
iva, R.
19, 5786, 1991
e 1 gene sequence from the primate *Saguinus imperator* isolated with PCR
X: S22582; MUID:92051332; PMID:1840669
82
DNA
<QUE>
s: EMBL:X61678; NID:958405; PIDN:CAA43853.1; PID:94494091
rs translated the codon TAC for residue 43 as Thr
erm histone
osomal protein; DNA binding; nucleus; spermatogenesis
2.5%; Score 7; DB 2; Length 50;
larity 100.0%; Pred. No. 28;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GRRG 48
|||||
GRRG 24
American opossum
is virginiana; *Didelphis marsupialis virginiana* (North American opossum)
95 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
15
Nishikawa, S.; Connor, W.; Dixon, G.H.
215, 63-72, 1993
ization of a marsupial sperm protamine gene and its transcripts from
c: S34045; MUID:93345500; PMID:8344286
15
lary
DNA
<WIN>
s: EMBL:X74044; NID:9407062; PIDN:CAA52193.1; PID:9407063
erm histone
inding; nucleus
2.5%; Score 7; DB 2; Length 58;
larity 100.0%; Pred. No. 32;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
GRRG 49
|||||
GRRG 41

RESULT 23
F87604
hypothetical protein CC2870 [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: F87604
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: F87604
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <STO>
A:Cross-references: GB:AE005673; NID:gl3424486; PIDN:AAK24834.1; GSFI
C:Genetics:
A:Gene: CC2870
Query Match 2.5%; Score 7; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 59 LALGLGL 65
Db 26 LALGLGL 32
|||||
RESULT 24
F87993
protein ZC334.3 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: F87993
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode *C. elegans*: a platform for i
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.a
A>Note: published errata appeared in Science 283, 35, 1999; Science 2
A:Accession: F87993
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <STO>
A:Cross-references: GB:chr I; PIDN:CAB04964.1; PID:g3881432; GSPDB:GN
A>Note: predicted using Genefinder
C:Genetics:
A:Gene: ZC334.3
A:Map position: 1
Query Match 2.5%; Score 7; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 55 LLVPLAL 61
Db 33 LLVPLAL 39
|||||
RESULT 25
PH1560
lg heavy chain V region (clone VH32) - human (fragment)
C:Species: *Homo sapiens* (man)
C>Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 21-Jan-1994
C:Accession: PH1560
R:Rassenti, L.Z.; Kipps, T.J.
J. Exp. Med. 177, 1039-1046, 1993
A>Title: Lack of extensive mutations in the VH5 genes used in common i
A:Reference number: PH1557; MUID:93210459; PMID:7681468
A:Accession: PH1560

06:25:20 2004

us-09-245-198a-4.oligo.rpr

abditis elegans
#sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
QY 65 LALACLG 71
DB 9 LALACLG 15
RESULTS 33
T37116
probable transposase, truncated [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Sep
C:Accession: T37116
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21588
A:Accession: T37116
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-146 <SAU>
A:Cross-references: EMBL:AL109950; PIDN:CAB52967.1; GSPDB:GN00070; SC
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SC08DB:SCJ4.33c
C:Superfamily: Synchocystis transposase sll1710
Query Match 2.5%; Score 7; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 70 LGLLAV 76
DB 102 LGLLAV 108
RESULTS 34
A71217
hypothetical protein PH2001 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun
C:Accession: A71217
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.;
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 5, 55-76, 1998
A>Title: Complete sequence and gene organization of the genome of a hy
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71217
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-147 <KAW>
A:Cross-references: GB:AP000007; GB:AP000001; NID:g3236134; NID:g32361
A:Experimental source: strain OT3
A>Note: this accession replaces an interim accession for a sequence re
A>Note: this sequence is split into two separate translations in GenBa
C:Genetics:
A:Gene: PH2001
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH2001
Query Match 2.5%; Score 7; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 67 LALICGL 73
DB 45 LALICGL 51
RESULTS 35
T08734
hypothetical protein DKEP566F0546.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug
C:Accession: T08734

abditis elegans
#sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
QY 65 LALACLG 71
DB 9 LALACLG 15
RESULTS 33
T37116
probable transposase, truncated [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Sep
C:Accession: T37116
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21588
A:Accession: T37116
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-146 <SAU>
A:Cross-references: EMBL:AL109950; PIDN:CAB52967.1; GSPDB:GN00070; SC
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SC08DB:SCJ4.33c
C:Superfamily: Synchocystis transposase sll1710
Query Match 2.5%; Score 7; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 70 LGLLAV 76
DB 102 LGLLAV 108
RESULTS 34
A71217
hypothetical protein PH2001 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun
C:Accession: A71217
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.;
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 5, 55-76, 1998
A>Title: Complete sequence and gene organization of the genome of a hy
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71217
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-147 <KAW>
A:Cross-references: GB:AP000007; GB:AP000001; NID:g3236134; NID:g32361
A:Experimental source: strain OT3
A>Note: this accession replaces an interim accession for a sequence re
A>Note: this sequence is split into two separate translations in GenBa
C:Genetics:
A:Gene: PH2001
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH2001
Query Match 2.5%; Score 7; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 67 LALICGL 73
DB 45 LALICGL 51
RESULTS 35
T08734
hypothetical protein DKEP566F0546.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug
C:Accession: T08734

```

; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
; Protein Sequence Database, May 1999
; Z16474
; RNA
; <OTT>
; EMBL:AL050075
; rice: fetal kidney; clone DKFZp566F0546
546.1
      2.5%; Score 7; DB 2; Length 150;
arity 100.0%; Pred. No. 72;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
LGP 240
|||
LGP 38

- rice
ativa (rice)
} #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
}
; H.; Yamada, T.; Tanaka, K.; Takeuchi, S.; Nakamura, R.; Matsuda, T.
; 239-248, 1993
; ture and expression of rice seed allergenic proteins belonging to the
; : S31078; MUID:93144699; PMID:7678765
; RNA
; ADA>
; EMBL:D11430; NID:g218196; PIDN:BAA01996.1; PID:g218197
; t alpha-amylase inhibitor

      2.5%; Score 7; DB 2; Length 157;
arity 100.0%; Pred. No. 75;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
VS 78
|||
VS 17

tiva (rice)
; #sequence_revision 24-Mar-1999 #text_change 16-Jul-1999
; J.H.; Eun, M.Y.
; BL Data Library, January 1998
; eotide sequence of rice allergenic protein.
; Z14691
; Y; translated from GB/EMBL/DBJ
; NA
; YUN>
; EMBL:AF042200; NID:g2827315; PIDN:AAB99797.1; PID:g2827316
; ce: strain Nipponbare
; t alpha-amylase inhibitor

      2.5%; Score 7; DB 2; Length 157;
arity 100.0%; Pred. No. 75;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
VS 78
|||
VS 17

```

```

RESULT 38
A75567
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar
C:Accession: A75567
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.
; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus r
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75567
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-157 <WHI>
A:Cross-references: GB:AE001867; GB:AE000513; NID:g6457693; PIDN:AAF09
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0033
A:Map position: 1

Query Match      2.5%; Score 7; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

QY      261 IRTLPPWA 267
DB      22 IRTLPPWA 28

RESULT 39
E75530
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar
C:Accession: E75530
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.
; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus r
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75530
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-157 <WHI>
A:Cross-references: GB:AE001895; GB:AE000513; NID:g6458024; PIDN:AAF09;
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0352
A:Map position: 1

Query Match      2.5%; Score 7; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

QY      43 RRRGRGG 49
DB      144 RRRGRGG 150

RESULT 40
S59925
allergen RASB precursor - rice
C:Species: Oryza sativa (rice)
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-
C:Accession: S59925
R:Alvarez, A.M.; Adachi, T.; Nakase, M.; Aoki, N.; Nakamura, R.; Matsuc
; Biochim. Biophys. Acta 1251, 201-204, 1995
A>Title: Classification of rice allergenic protein cDNAs belonging to t
A:Reference number: S59922; MUID:95399441; PMID:7669811
A:Accession: S59925

```


1ary
 RNA
 <ALV>
 : EMBL:D42142; NID:gl398917; PIDN:BAA07713.1; PID:gl398918
 at alpha-amylase inhibitor
 2.5%; Score 7; DB 2; Length 160;
 larity 100.0%; Pred. No. 76;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 AVVS 78
 |||||
 AVVS 17
 protein hlpA precursor - Escherichia coli (strain K-12)
 : DNA-binding 17K protein; histone-like protein hlp
 ichia coli
 19 #sequence_revision 30-Jun-1989 #text_change 01-Mar-2002
 4; A38063; S13728; B64742; I54944; S20426
 e, K.
 1988
 and sequencing of the gene for the DNA-binding 17K protein of Escherich
 : JT0304; MUID:88329735; PMID:2843433
 14
 <HOL>
 : GB:M21118; NID:gl47821; PIDN:AAA24630.1; PID:gl47822
 13
 protein
 <HO2>
 : GB:M21118; NID:gl47821; PIDN:AAA24630.1; PID:gl47822
 : S13728; MUID:91100302; PMID:1987124
 18
 1ary
 <DIC>
 : EMBL:X54797; NID:941468; PIDN:CAA38567.1; PID:941469
 : strain K-12, substrain MG1655
 Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 lau, B.; Shao, Y.
 1462, 1997
 ete genome sequence of Escherichia coli K-12.
 : A64720; MUID:97426617; PMID:9278503
 2
 acid sequence not shown; translation not shown
 NA
 <BLAT>
 : GB:AE000127; GB:U00096; NID:gl786370; PIDN:AAC73289.1; PID:gl786375;
 : strain K-12, substrain MG1655
 i, P.; Vaara, M.
 1223-1229, 1991
 Gene of Yersinia enterocolitica: cloning, sequencing, expression, and
 : I54944; MUID:91123198; PMID:1991717
 4
 ed from GB/EMBL/DBJ
 NA
 L', 16-148, E', 150-152, I', 154-161 <RES>
 : EMBL:X75465; NID:9432661; PIDN:CAA53207.1; PID:9432662
 A protein has been believed to be a histone-like constituent of bacter
 1-binding 17K protein
 ine protein
 inal sequence #status predicted <SIG>
 outer membrane protein hlpA #status predicted <MAT>
 2.5%; Score 7; DB 1; Length 161;

Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 QY 62 GLGLALA 68
 DB 9 GLGLALA 15
 RESULT 42
 D90651
 histone-like protein hlpA [imported] - Escherichia coli (strain O157:
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-
 C:Accession: D90651
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yoko
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shina
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia co
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: D90651
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-161 <HAV>
 A:Cross-references: GB:BA000007; PIDN:BA33603.1; PID:gl3359636; GSPDI
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC80180
 C:Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 62 GLGLALA 68
 DB 9 GLGLALA 15
 RESULT 43
 D85502
 hypothetical protein hlpA [imported] - Escherichia coli (strain O157:
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep
 C:Accession: D85502
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.;
 iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potam
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: D85502
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-161 <STO>
 A:Cross-references: GB:AB005174; NID:gl2512906; PIDN:AAG54480.1; GSPDI
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: hlpA
 C:Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 62 GLGLALA 68
 DB 9 GLGLALA 15
 RESULT 44
 T24937
 hypothetical protein W03C9.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct

```

7: T26123
WBL Data Library, July 1995
: Z19957
7
ary: translated from GB/EMBL/DBJ
NA
<W11>
: EMBL:Z50015; PIDN:CAA90314.1; GSPDB:GN00020; CESP:W03C9.4
rce: clone T15G9
scough, R.
WBL Data Library, October 1995
: Z20155
3
ary: translated from GB/EMBL/DBJ
NA
<W12>
: EMBL:Z66516; PIDN:CAA91361.1; GSPDB:GN00020; CESP:W03C9.4
rce: clone W03C9
.4
4/1; 120/3; 128/1
2.5%; Score 7; DB 2; Length 162;
arity 100.0%; Pred. No. 77;
conservative 0; Mismatches 0; Indels 0; Gaps 0;
ASS 237
|||
ASS 118
|||

in 424 - Spingomonas aromaticivorans plasmid pNL1
monas aromaticivorans
) #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
)
llwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; SENSEN, C.W.; G
WBL Data Library, July 1998
plete sequence of a 184 kb catabolic plasmid from Spingomonas aromati
: Z20992
)
ary: translated from GB/EMBL/DBJ
NA
:ROM>
: EMBL:AF079317; NID:g3378261; PID:g3378314; PIDN:AA003897.1
)NL1
ngomonas aromaticivorans hypothetical protein 424
2.5%; Score 7; DB 2; Length 162;
arity 100.0%; Pred. No. 77;
conservative 0; Mismatches 0; Indels 0; Gaps 0;
AA 144
|||
AA 34

n - Synecocystis sp. (strain PCC 6803)
ystis sp.
) #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
)
S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
, 1996
analysis of the genome of the unicellular cyanobacterium Synecocystis

```

5.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76920
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-164 <KAN>
A;Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BAALF
A;Note: the nucleotide sequence was submitted to the EMBL Data Library

Query Match 2.5%; Score 7; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

Qy 104 SELNPOT 110
|||
Db 37 SELNPOT 43

RESULT 47
S44789
D2007.4 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar
C;Accession: S44789
R;Favell, A.D.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid D2007.
A;Reference number: S44619
A;Accession: S44789
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-170 <FAV>
A;Cross-references: EMBL:L16560; NID:g289666; PID:g289670
C;Genetics:
A;Introns: 43/2; 121/3

Query Match 2.5%; Score 7; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

Qy 221 VLALRCL 227
|||
Db 94 VLALRCL 100

RESULT 48
D87638
transcription regulator, GntR family [imported] - Caulobacter crescent
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr
C;Accession: D87638
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.;
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-174 <STO>
A;Cross-references: GB:AE005673; NID:g13424808; PIDN:AAK25104.1; GSPDB
C;Genetics:
A;Gene: CC3142

Query Match 2.5%; Score 7; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G

Qy 138 ARRATAA 144
|||
Db 151 ARRATAA 157

```

n precursor - goat
aegagrus hircus (domestic goat)
85 #sequence revision 12-Apr-1996 #text_change 22-Jun-1999
20: S14507; S42800; S42801
unitzer, G.; Schrank, B.; Stangl, A.
Physiol. Chem. 360, 1595-1604, 1979
o acid sequence of goat beta-lactoglobulin.
r: A91682; MUID:80070611; PMID:511095
20
protein
0 <PRE>
11, A.; Sanchez, A.
EMBL Data Library, March 1991
r: S14507
07
mRNA
<POL>
s: EMBL:X58471; NID:g967; PIDN:CAA41385.1; PID:g968
EMBL Data Library, January 1993
r: S42800
00
mRNA
<KIM>
s: EMBL:Z19569; NID:g437751; PIDN:CAA79623.1; PID:g437752
01
mRNA
<K12>
3: EMBL:Z19570; NID:g437753; PIDN:CAA79624.1; PID:g437754
physiological conditions beta-lactoglobulin exists as an equilibrium mi
pocalin; lipocalin homology
gnal sequence #status predicted <SIG>
beta-lactoglobulin #status predicted <MAT>
lipocalin homology <LIP>
Disulfide bonds: #status predicted
2.5%; Score 7; DB 1; Length 180;
arity 100.0%; Pred. No. 84;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ALAC 69
|||||
ALAC 14
1 precursor - sheep
: beta-lactoglobulin A; beta-lactoglobulin B; beta-lactoglobulin C; bet
:entalis aries, Ovis ammon aries (domestic sheep)
35 #sequence revision 19-Apr-1996 #text_change 22-Jun-1999
19: JQ0748; A30011; B30011; S02136; A03221; S04955
ghan, M.; Simons, J.P.; Clark, A.J.
1990
isation of the alleles encoding ovine beta-lactoglobulins A and B.
r: JQ0748; MUID:91007276; PMID:1976573
19
DNA
<ALIB>
s: GB:M32232
rce: beta-lactoglobulin B
18
DNA
Y', 39-180 <ALIA>
s: GB:M32232
rce: beta-lactoglobulin A
A.J.
415-426, 1988
ization of the gene encoding ovine beta-lactoglobulin. Similarity to d

```

A;Reference number: A92942; MUID:88172489; PMID:3351935
A;Accession: A30011
A;Molecule type: DNA
A;Residues: 1-180 <ALII>
A;Cross-references: GB:X14971
A;Experimental source: beta-lactoglobulin I
A;Accession: B30011
A;Molecule type: DNA
A;Residues: 1-37, 'Y', 39-102, 'N', 104-180 <ALI2>
A;Cross-references: GB:X07009
A;Experimental source: beta-lactoglobulin II
R;Harris, S.; Ali, S.; Anderson, S.; Archibald, A.L.; Clark, A.J.
Nucleic Acids Res. 16, 10379-10380, 1988
A;Title: Complete nucleotide sequence of the genomic ovine beta-lacto
A;Reference number: S02136; MUID:89057492; PMID:3194215
A;Accession: S02136
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-180 <HAR>
A;Cross-references: EMBL:X12817; NID:g1313; PIDN:CAA31305.1; PID:g131
R;Gaye, P.; Hue-Delahaie, D.; Mercier, J.C.; Soulier, S.; Vilotte, J.
Biochimie 68, 1097-1107, 1986
A;Title: Ovine beta-lactoglobulin messenger RNA: nucleotide sequence
A;Reference number: A25136; MUID:87049827; PMID:3096387
A;Accession: A25136
A;Molecule type: mRNA
A;Residues: 1-180 <GAY>
A;Cross-references: GB:X04520; NID:g1315; PIDN:CAA28204.1; PID:g1316
R;Preaux, G.; Braunitzer, G.; Kolde, H.J.
Arch. Int. Physiol. Biochim. 88, B45-B46, 1980
A;Title: Primary structure of ovine beta-lactoglobulin.
A;Reference number: A03221; MUID:80219294; PMID:6155855
A;Accession: A03221
A;Molecule type: protein
A;Residues: 19-37, 'Y', 39-180 <PRE>
R;Brhardt, G.; Godovac-Zimmermann, J.; Conti, A.
Biol. Chem. Hoppe-Seyler 370, 757-762, 1989
A;Title: Isolation and complete primary sequence of a new ovine wild-
A;Reference number: S04955; MUID:89374823; PMID:2775495
A;Accession: S04955
A;Molecule type: protein
A;Residues: 19-37, 'Y', 39-165, 'Q', 167-180 <ERH>
A;Experimental source: beta-lactoglobulin C
C;Comment: This protein is the major milk whey protein of ruminants a
C;Comment: Under physiological conditions beta-lactoglobulin exists a
C;Genetics:
A;Gene: BLG
A;Introns: 32/3; 79/2; 104/1; 141/1; 176/1
C;Superfamily: lipocalin; lipocalin homology
C;Keywords: milk; polymorphism
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-180/Product: beta-lactoglobulin #status experimental <MAT>
F;28-178/Domain: lipocalin homology <LIP>
F;84-178, 124-137/Disulfide bonds: #status predicted
Query Match 2.5%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 63 LGLALAC 69
DB 8 LGLALAC 14
RESULT 51
B60738
insulin-like growth factor II precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Apr-1993 #sequence_revision 30-Sep-1993 #text_change 13-Nov
C;Accession: S12614; B60738
R;Catchpole, I.R.; Engstroem, W.
Nucleic Acids Res. 18, 6430, 1990
A;Title: Nucleotide sequence of a porcine insulin-like growth factor I

: SL2614; MUID:91057136; PMID:2243790
 4 RNA
 <CAT>
 wens, P.C.; McNeill, K.A.; Wallace, J.C.; Ballard, F.J.
 ion, amino acid sequences and assay cross-reactivities of porcine insu
 : A60738; MUID:90039035; PMID:2809477
 3
 rotein
 'X', 81-91 <FRA>
 alin
 al sequence #status predicted <SIG>
 insulin-like growth factor II #status experimental <MAT>
 arboxyl-terminal propeptide (E peptide) #status predicted <CTP>
 5/Disulfide bonds: #status predicted
 2.5%; Score 7; DB 2; Length 181;
 irity 100.0%; Pred. No. 85; Indels 0; Gaps 0;
 nservative 0; Mismatches 0;
 AL 61
 |||||
 AL 18
 |||||
 ke protein [imported] - Arabidopsis thaliana
 ois thaliana (mouse-ear cress)
 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 ropean Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 /, 1999
 nd analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A85001; MUID:20083488; PMID:10617198
 ry
 RA
 .STO>
 GB:NC_001368; NID:g7269845; PIDN:CAB79704.1; GSPDB:GN00140
 2.5%; Score 7; DB 2; Length 187;
 irity 100.0%; Pred. No. 87;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 VS 78
 |||||
 VS 24
 n C35D10.8 - Caenorhabditis elegans
 bditis elegans
 #sequence_revision 25-Apr-1997 #text_change 20-Sep-1999
 ough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton,
 A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johnston, L.; Jon
 M.; Parsons, J.; Percy, C.; Rifkin, L.; Roopra, A.; Saunders, D.
 BL Data Library, February 1995
 n, R.; Smaldon, N.; Smith, A.; Sonhammer, E.; Staden, R.; Sulston, J
 u, P.
 C. elegans genome project: Contiguous nucleotide sequence of over two
 S72566
 A
 WIL>
 EMBL:U21324; NID:g687879; PID:g687888
 ce: strain Bristol N2

C;Genetics:
 A;Map position: 3
 A;Introns: 42/2; 88/3; 126/3; 176/3
 A;Note: C35D10.8
 C;Superfamily: Caenorhabditis elegans hypothetical protein C35D10.8
 Query Match 2.5%; Score 7; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 QY 136 TRARRAI 142
 |||||
 Db 98 TRARRAI 104
 |||||
 RESULT 54
 AH0592
 conserved hypothetical protein ydgQ [imported] - Salmonella enterica s
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov
 C;Accession: AH0592
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; W
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.;
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton
 A;Title: Complete genome sequence of a multiple drug resistant Salmone
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AH0592
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-230 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD01913.1; PID:g16502755; GSPDB
 C;Genetics:
 A;Gene: ydgQ
 C;Superfamily: conserved hypothetical protein HI1688
 Query Match 2.5%; Score 7; DB 2; Length 230;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 QY 60 ALGLGLA 66
 |||||
 Db 38 ALGLGLA 44
 |||||
 RESULT 55
 B64920
 probable membrane protein ydgQ - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar
 C;Accession: B64920
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland,
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: B64920
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-231 <BLAT>
 A;Cross-references: GB:AE000258; GB:U00096; NID:g2367121; PIDN:AAC74704
 A;Experimental source: strain K-12, substrain MG1655
 C;Genetics:
 A;Gene: ydgQ
 A;Start codon: GTG
 C;Superfamily: conserved hypothetical protein HI1688
 C;Keywords: transmembrane protein
 F;18-34/Domain: transmembrane #status predicted <TM01>
 F;38-54/Domain: transmembrane #status predicted <TM02>
 F;71-87/Domain: transmembrane #status predicted <TM03>
 F;97-113/Domain: transmembrane #status predicted <TM04>

transmembrane #status predicted <TM05>
transmembrane #status predicted <TM06>

2.5%; Score 7; DB 2; Length 231;
Identity 100.0%; Pred. No. 1e+02;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LGLA 66
|||||
LGLA 44

ein ECs2341 [imported] - Escherichia coli (strain O157:H7, substrain R1)
ichia coli

01 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

21
kino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
unaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

, 2001
genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
r: A99629; MUID:21156231; PMID:11258796

21
nary
DNA
<HAY>

s: GB:BA000007; PIDN:BA035764.1; PID:gl3361808; GSPDB:GN00154
urce: strain O157:H7, substrain RMD 050952

nserved hypothetical protein HI1688

2.5%; Score 7; DB 2; Length 231;
Identity 100.0%; Pred. No. 1e+02;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LGLA 66
|||||
LGLA 44

ein ydgQ [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
ichia coli

01 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

70
unkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
rk, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamowski, K.; Apodaca,
13, 2001
sequence of enterohemorrhagic Escherichia coli O157:H7.

r: A85480; MUID:21074935; PMID:11206551

70
nary
DNA
<STO>

r: GB:AE005174; NID:gl2515621; PIDN:AAG56621.1; GSPDB:GN00145; UWGP:Z26
urce: strain O157:H7, substrain EDL933

nserved hypothetical protein HI1688

2.5%; Score 7; DB 2; Length 231;
Identity 100.0%; Pred. No. 1e+02;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LGLA 66
|||||
LGLA 44

C86665

amino acid ABC transporter permease protein [imported] - Lactococcus l
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-De
C:Accession: C86665

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Wei
Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium La
A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: C86665

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-231 <STO>

A;Cross-references: GB:AE005176; PID:gl2723189; PIDN:AAK04421.1; GSPD

A;Experimental source: strain IL1403

C:Genetics:

A;Gene: ydCC

C:Superfamily: ABC transporter permease protein

Query Match 2.5%; Score 7; DB 2; Length 231;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 56 LVPLALG 62

|||||

Db 107 LVPLALG 113

RESULT 59

S60767

ribonuclease III - Coxiella burnetii

C:Species: Coxiella burnetii

C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 22-Ju

C:Accession: S60767

R;Zuber, M.; Hoover, T.A.; Powell, B.S.; Court, D.L.

Mol. Microbiol. 14, 291-300, 1994

A;Title: Analysis of the rnc locus of Coxiella burnetii.

A;Reference number: S60767; MUID:95131751; PMID:7830573

A;Accession: S60767

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-233 <ZUB>

A;Cross-references: EMBL:L27436; NID:9439870; PIDN:AAA69690.1; PID:94

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat h

F,150-223/Domain: double-stranded RNA-binding repeat homology <DSR>

Query Match 2.5%; Score 7; DB 2; Length 233;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 10 ARRLPLP 16

|||||

Db 164 ARRLPLP 170

RESULT 60

AB0273

probable membrane protein YPO2240 [imported] - Yersinia pestis (strain
C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov

C:Accession: AB0273

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; White

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of pl

A;Reference number: AB00001; MUID:21470413; PMID:11586360

A;Accession: AB0273

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-233 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC91046.1; PID:gl5980240; GSPDE

C:Genetics:

erved hypothetical protein H11688
 2.5%; Score 7; DB 2; Length 233;
 rity 100.0%; Pred. No. 1.1e+02;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 LA 66
 ||
 LA 44
 synthase-like protein [imported] - Arabidopsis thaliana
 sis thaliana (mouse-ear cress)
 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
 ropan Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 , 1999
 nd analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A85001; MUID:20083488; PMID:10617198
 ry
 A
 STO>
 GB_NC_001268; NID:g7267660; PIDN:CAB78088.1; GSPDB:GN00140
 ransporting ATP synthase delta chain
 2.5%; Score 7; DB 2; Length 234;
 rity 100.0%; Pred. No. 1.1e+02;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 SS 237
 ||
 SS 54
 nslocating NADH dehydrogenase (ubiquinone) (EC 1.6.5.-) nqrD chain H1
 lus influenzae
 #sequence_revision 18-Aug-1995 #text_change 21-Jul-2000
 Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
 ott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
 C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
 2, 1995
 .L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 me random sequencing and assembly of Haemophilus influenzae Rd.
 A64000; MUID:95350630; PMID:7542800
 acid sequence not shown; translation not shown
 IA
 TIGR>
 GB:U32841; GB:L42023; NID:gl574529; PIDN:AAC23334.1; PID:gl574540; T
 erved hypothetical protein H11688
 ductase
 2.5%; Score 7; DB 2; Length 235;
 rity 100.0%; Pred. No. 1.1e+02;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 ILA 66
 ||
 ILA 68

probable cytochrome b-561 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-
 C:Accession: A86387
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; W
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Kha
 A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mai
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.;
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A86387
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-236 <STO>
 A:Cross-references: GB:A8005172; NID:gil079498; PIDN:AAG29209.1; GSPDB:
 C:Genetics:
 A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 236;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps
 QY 61 LGLGLAL 67
 Db 194 LGLGLAL 200

RESULT 64

H70866
 hypothetical protein Rv2473 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-
 C:Accession: H70866
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Har
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squ
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the
 A:Reference number: A70500; MUID:98295987; PMID:96344230
 A:Accession: H70866
 A>Status: preliminary; nucleic acid sequence not shown; translation not
 A:Molecule type: DNA
 A:Residues: 1-238 <COL>
 A:Cross-references: GB:AL021246; GB:AL123456; NID:g3261507; PIDN:CAA161
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv2473

Query Match 2.5%; Score 7; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 74 LAVVSLG 80
 Db 90 LAVVSLG 96

RESULT 65

G83208
 conserved hypothetical protein PA3494 [imported] - Pseudomonas aerugin
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 27-Nov-
 C:Accession: G83208
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.
 Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000

genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
AB2950; MUID:20437337; PMID:10984043

Query Match 2.5%; Score 7; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
A: Map position: 5
A: Introns: 41/1; 84/2; 144/3; 195/3

served hypothetical protein H1688

Query Match 2.5%; Score 7; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GLA 66

GLA 46

synthesis protein moeB - Haemophilus influenzae (strain Rd KW20)

plus influenzae
5 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
4
; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
cott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
C.; Fine, L.D.; Fritschman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
12, 1995
C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
ome random sequencing and assembly of Haemophilus influenzae Rd.
; A64000; MUID:95350630; PMID:7542800

acid sequence not shown; translation not shown

NA

<TIGR>
; GB:U32823; GB:L42023; NID:G1574281; PIDN:AAC23099.1; PID:G1574288;

pterin biosynthesis
yldopterin biosynthesis protein moeB
ienum; molybdopterin biosynthesis

Query Match 2.5%; Score 7; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

NRQ1 185

NRQ1 16

ein ZC196.8 - *Caenorhabditis elegans*
habditis elegans

99 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
42

EMBL Data Library, April 1997

e sequence of *C. elegans* cosmid ZC196.

r: Z20115

nary; translated from GB/EMBL/DBDJ

DNA

<MUR>

s: EMBL:U97007; PIDN:AAB52298.1; GSPDB:GN00023; CESP:ZC196.8

source: strain Bristol N2; clone ZC196

6.8

Query Match 2.5%; Score 7; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AASSLGP 240

Db 225 AASSLGP 231

RESULT 68

A46066
lymphotoxin beta - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul
C:Accession: A46066
R:Browning, J.L.; Ngam-ek, A.; Lawton, P.; DeMarinis, J.; Tizard, R.;
Cell 72, 847-856, 1993
A:Title: Lymphotoxin beta, a novel member of the TNF family that forms
A:Reference number: A46066; MUID:93208881; PMID:7916655
A:Accession: A46066
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-244

A:CROSS-references: GB:L11015; NID:G92276; PIDN:AAA36191.1; PID:G2922
A:Note: sequence extracted from NCBI backbone (NCBIN:128086, NCBI:128
C:Keywords: transmembrane protein

Query Match 2.5%; Score 7; DB 2; Length 244;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GLYLYLC 200

Db 131 GLYLYLC 137

RESULT 69

B32352
molybdopterin biosynthesis protein moeB - *Escherichia coli* (strain K-
N:Alternate names: molybdopterin-converting factor chlN
C:Species: *Escherichia coli*
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 01-Ma
C:Accession: B32352; B64820
R:Nohno, T.; Kasai, Y.; Saito, T.
J. Bacteriol. 170, 4097-4102, 1988
A:Title: Cloning and sequencing of the *Escherichia coli* chlN operon
A:Reference number: A32352; MUID:8814906; PMID:3045084
A:Accession: B32352
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-249 <NOH>
A:CROSS-references: GB:M21151; NID:G145538; PIDN:AAA23580.1; PID:G145
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64820
A:Status: preliminary; nucleic acid sequence not shown; translation n
A:Molecule type: DNA
A:Residues: 1-249 <BLAT>
A:CROSS-references: GB:AE000185; GB:U00096; NID:G1787047; PIDN:AACT739
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: moeB; chlN
A:Map position: 18 min

C:Function:
A:Pathway: molybdopterin biosynthesis
C:Superfamily: molybdopterin biosynthesis protein moeB
C:Keywords: molybdenum; molybdopterin biosynthesis

2.5%; Score 7; DB 2; Length 249;
 100.0%; Pred. No. 1.1e+02;
 Mismatches 0; Indels 0; Gaps 0;
 QOI 185
 ||||
 QOI 16

ynthesis MoeB protein [imported] - Escherichia coli (strain O157:H7, s
 chia coli
 1 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 1
 ino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 naga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 2001
 genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
 : A99629; MUID:21156231; PMID:11258796
 1
 ary
 NA
 <HAY>
 : GB:BA000007; PIDN:BA34327.1; PID:gl3360363; GSPDB:GN00154
 rce: strain O157:H7, substrain RMD 0509952

ybdopterin biosynthesis protein moeB
 2.5%; Score 7; DB 2; Length 249;
 100.0%; Pred. No. 1.1e+02;
 Mismatches 0; Indels 0; Gaps 0;
 QOI 185
 ||||
 QOI 16

ynthesis [imported] - Escherichia coli (strain O157:H7, substrain EDL9
 ichia coli
 01 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 32
 ck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
 33, 2001
 sequence of enterohemorrhagic Escherichia coli O157:H7.
 r: A85480; MUID:21074935; PMID:11206551
 92
 nary
 DNA
 <STO>
 s: GB:AE005174; MID:gl2513829; PIDN:BA55199.1; GSPDB:GN00145; UWGP:Z10
 urce: strain O157:H7, substrain EDL933

lybdopterin biosynthesis protein moeB
 2.5%; Score 7; DB 2; Length 249;
 100.0%; Pred. No. 1.1e+02;
 Mismatches 0; Indels 0; Gaps 0;
 QOI 185
 ||||
 QOI 16

ynthesis MoeB protein [imported] - Salmonella enterica subsp. enterica

C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov
 C;Accession: AI0602
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; W
 Ch, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.;
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton
 A;Title: Complete genome sequence of a multiple drug resistant Salmone
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AI0602
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-249 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD05291.1; PID:gl502055; GSPDE
 C;Genetics:
 A;Gene: STY0884
 C;Superfamily: molybdopterin biosynthesis protein moeB

Query Match 2.5%; Score 7; DB 2; Length 249;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 QY 179 LRYNRQI 185
 |||||
 Db 10 LRYNRQI 16

RESULT 73
 AI3154
 1821 family transposase istB [imported] - Agrobacterium tumefaciens (;
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-No
 C;Accession: AT3154
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, J
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tum
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AI3154
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-252 <KUR>
 A;Cross-references: GB:AE008689; PIDN:AAL45655.1; PID:gl7743380; GSPD
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: istB
 A;Map position: linear chromosome
 C;Superfamily: DNA replication protein dnaC

Query Match 2.5%; Score 7; DB 2; Length 252;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 QY 60 ALGLGLA 66
 |||||
 Db 120 ALGLGLA 126

RESULT 74
 H98132
 istB protein (AJ238712) [imported] - Agrobacterium tumefaciens (strai
 C;Species: Agrobacterium tumefaciens
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nc
 C;Accession: H98132
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qi
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lapp
 Science 294, 2323-2328, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Age

c: A97359; MUID:21608551; PMID:11743194
32
ary
DNA
<KUR>
s: GB:AE007870; PIDN:AAK88586.1; PID:gl5158297; GSPDB:GN00170

linear chromosome
A replication protein dnaC
2.5%; Score 7; DB 2; Length 252;
larity 100.0%; Pred.No.1.1e+02;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LGLA 66
||||
LGLA 126

synthesis MoeB protein VCA0618 [imported] - Vibrio cholerae (strain N16
cholerae
00 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
39
; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
olaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.
s, J.G.; Venter, J.C.; Fraser, C.M.
83, 2000
ence of both chromosomes of the cholera pathogen Vibrio cholerae.
r: A82035; MUID:20406833; PMID:10952301
39
nary
DNA
<HEI>
s: GB:AE004391; GB:AE003853; NID:g9658015; PIDN:AAF96519.1; GSPDB:GN001
urce: serogroup O1; strain N16961; biotype El Tor

lydopterin biosynthesis protein moeB
2.5%; Score 7; DB 2; Length 254;
larity 100.0%; Pred.No.1.1e+02;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

NRQI 185
||||
NRQI 18

April 7, 2004, 17:59:25
s

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.
 n search, using sw model
 fil 7, 2004, 17:48:13 ; Search time 17 Seconds
 (without alignments)
 869.878 Million cell updates/sec
 -09-245-198a-4
 i
 MSLLDFEISARRLPLRSLG.....PWAHLXAAPLTYFGLFQVH 284

GO
 pop 60.0 , Gapext 60.0

[681 seqs, 52070155 residues

s satisfying chosen parameters: 141681

hth: 0

hth: 2000000000

string first 100 summaries

visProt_42:*

the number of results predicted by chance to have a
 r than or equal to the score of the result being printed,
 ad by analysis of the total score distribution.

SUMMARIES

seq	ch	Length	DB	ID	Description
7.7	249	1	TN12	HUMAN	O43508 homo sapien
1.3	225	1	TN12	MOUSE	O54907 mus musculu
3.5	111	1	YFBW	ECOLI	Q47377 escherichia
3.2	733	1	PSAB	ODOSI	P49480 odontellia s
3.2	734	1	PSAB	CYACA	Q9t1g6 cyanidium c
2.8	179	1	ADHS	GLUOX	O05544 gluconobact
2.8	220	1	Y304	BRUME	O8vd73 brucella me
2.8	220	1	Y3J1	BRUSU	O8fv59 brucella su
2.8	317	1	MSHR	PANTR	Q9trk4 pan troglod
2.8	379	1	FDHB	METJA	Q60316 methanococc
2.8	422	1	ZP3	MESAU	P23491 mesocricetu
2.8	576	1	CVDC	HABIN	P45081 haemophilus
2.8	586	1	GGT5	HUMAN	P36269 homo sapien
2.5	24	1	HSP3	OCTVU	P83215 octopus vul
2.5	30	1	HSP5	OCTVU	P83217 octopus vul
2.5	49	1	HSP1	SAGIM	P24714 saginus im
2.5	57	1	HSP1	DIDMA	P35305 didelphis m
2.5	115	1	A62F	DROME	O46202 drosophila
2.5	118	1	Y151	ECOLI	P03838 escherichia
2.5	131	1	IL13	MOUSE	P20109 mus musculu
2.5	131	1	IL13	RAT	P42203 rattus norv
2.5	147	1	YK01	PYRHO	O57781 pyrococcus
2.5	150	1	TNFC	PIG	Q9tsv8 sus scrofa
2.5	157	1	RA05	ORYSA	O01881 oryza sativ
2.5	161	1	HLPA	ECOLI	P11457 escherichia
2.5	170	1	YIM4	CABEL	P34378 caenorhabdi
2.5	180	1	LACB	BUBBU	P02755 bubalus bub
2.5	180	1	LACB	CAPHI	P02756 capra hircu
2.5	181	1	LACB	SHEEP	P02757 ovis aries
2.5	181	1	APT	SHEON	O8efg1 shewanella
2.5	217	1	DEFI	BIFLO	Q8g534 bifidobacte
2.5	230	1	RNFE	SALTY	Q8xex9 salmonella
2.5	231	1	RNFE	ECOL	P58344 escherichia

34	7	2.5	231	1	RNFE	ECOLI	P77179 esch
35	7	2.5	233	1	RNC	COXBU	P51837 coxi
36	7	2.5	233	1	RNFE	YERPE	O8zed4 yers
37	7	2.5	235	1	RNFE	HABIN	O57020 haem
38	7	2.5	239	1	TN14	MOUSE	Q9qyh9 mus
39	7	2.5	240	1	RNFE	PSEAE	Q9hyh5 pseu
40	7	2.5	243	1	MOEB	HABIN	F45211 haem
41	7	2.5	244	1	TNFC	HUMAN	Q06643 homo
42	7	2.5	244	1	TNFC	PANTR	Q86227 pan
43	7	2.5	249	1	LFTF	XANCP	Q8p996 xant
44	7	2.5	249	1	MOEB	ECOLI	P12282 esch
45	7	2.5	249	1	MOEB	SALTY	Q56087 salm
46	7	2.5	257	1	KDIX	SERMA	Q54435 serr
47	7	2.5	277	1	CN09	HUMAN	Q86t03 homo
48	7	2.5	310	1	TNFC	MARMO	Q9jml0 mar
49	7	2.5	316	1	ISPH	XANCP	Q8p64 xant
50	7	2.5	316	1	ISPH	XYLFA	Q9pas9 xyle
51	7	2.5	324	1	ODPB	BAGSU	P21882 baci
52	7	2.5	329	1	SRA6	CABEL	Q09208 caen
53	7	2.5	335	1	IAG2	RAT	O35777 ratt
54	7	2.5	344	1	LEU3	THEAQ	P24098 ther
55	7	2.5	352	1	LEU3	DEIRA	Q9rth9 dein
56	7	2.5	357	1	G6PT	MOUSE	P35576 mus
57	7	2.5	357	1	G6PT	RAT	P43428 ratt
58	7	2.5	358	1	PONT	RABIT	P27170 oryc
59	7	2.5	365	1	NQ08	THETH	Q60019 ther
60	7	2.5	394	1	BENE	ACICA	P07775 acin
61	7	2.5	396	1	DH11	XENLA	Q91610 xeno
62	7	2.5	398	1	DH12	XENLA	Q91611 xeno
63	7	2.5	402	1	SELP	BOVIN	P49907 bos
64	7	2.5	412	1	PGKP	ALCEU	P50320 alca
65	7	2.5	413	1	PGKC	ALCEU	P50319 alca
66	7	2.5	416	1	NH59	CABEL	Q9txj1 caen
67	7	2.5	418	1	CP16	RAT	P09006 ratt
68	7	2.5	419	1	ENO	PYRAE	Q8zve7 pyro
69	7	2.5	419	1	PGK	RALSO	Q8ylw6 rals
70	7	2.5	423	1	YJ54	YEAST	P47130 sacc
71	7	2.5	424	1	ZP3	MOUSE	P10761 mus
72	7	2.5	428	1	SYH	CHLMU	Q9pjj9 chia
73	7	2.5	455	1	PHR	STRGR	P12768 stre
74	7	2.5	461	1	PUCG	RHOCA	P23462 rhod
75	7	2.5	461	1	Y608	HABIN	Q57486 haem
76	7	2.5	464	1	SOX8	MOUSE	Q04886 mus
77	7	2.5	467	1	D4DR	HUMAN	P21917 homo
78	7	2.5	483	1	GLME	CLOCC	P80077 clos
79	7	2.5	485	1	GLME	CLOTT	Q05509 clos
80	7	2.5	505	1	PDI	HUMIN	P55059 humi
81	7	2.5	512	1	LNT	ECOLI	P23910 esch
82	7	2.5	512	1	LNT	SALTY	O87576 balm
83	7	2.5	518	1	ASB3	HUMAN	Q9y575 homo
84	7	2.5	525	1	ASB3	MOUSE	Q9wv72 mus
85	7	2.5	547	1	RM56	HUMAN	P83111 homo
86	7	2.5	551	1	RM56	MOUSE	Q9ep89 mus
87	7	2.5	600	1	S133	MOUSE	Q9ly63 mus
88	7	2.5	600	1	S133	RAT	Q9z0z5 ratt
89	7	2.5	653	1	MUTL	VIBCH	Q9kv13 vibr
90	7	2.5	720	1	PPTA	PSEAE	F45212 pseo
91	7	2.5	758	1	CSTA	MYCTU	P95095 myco
92	7	2.5	760	1	PO21	XENLA	P16143 xeno
93	7	2.5	810	1	NELI	HUMAN	Q92832 homo
94	7	2.5	810	1	NELI	RAT	Q62919 ratt
95	7	2.5	854	1	KDPD	RATRA	O34971 ratt
96	7	2.5	953	1	B3A4	RAT	Q8k4v2 ratt
97	7	2.5	955	1	B3A4	RABIT	Q9gkv1 oryc
98	7	2.5	983	1	B3A4	HUMAN	Q96q31 oryc
99	7	2.5	1072	1	ITAC	CHICK	P26007 gall
100	7	2.5	1076	1	IF3A	CABEL	P34339 caen

ALIGNMENTS

RESULT 1

STANDARD; PRT; 249 AA.

Rel. 41, Created)

Rel. 41, Last sequence update)

Rel. 42, Last annotation update)

s factor ligand superfamily member 12 (TNF-related weak apoptosis) (TWEAK) (APO3 ligand).

03L OR DR3LG.

(Human)

Chordata; Craniata; Vertebrata; Euteleostomi;

Primates; Catarrhini; Hominidae; Homo.

06;

1 N.A., AND N-TERMINUS OF SOLUBLE FORM.

liver, and Tonsil;

1415; PubMed=9405449;

ie Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,

Barcia I., Browning J.L.;

; secreted ligand in the tumor necrosis factor family that is apoptosis.";

1. 272:32401-32410(1997).

1 N.A.

Kidney;

1355; PubMed=9560343;

, Sheridan J.P., Pitti R.M., Brush J., Goddard A.,

on of a ligand for the death-domain-containing receptor

1:525-528(1998).

1 N.A.

1257; PubMed=12477932;

L., Feigold E.A., Grouse L.H., Derge J.G.,

, Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

, Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

, Marusina K., Farmer A., Rubin G.M., Hong L.,

, Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

Quellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

, Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

ton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

, Touchman J.W., Green E.D., Dickson M.C.,

, Grimwood J., Schmutz J., Myers R.M.,

, S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

Schein J.E., Jones S.J.M., Marra M.A.;

and initial analysis of more than 15,000 full-length

use cDNA sequences.";

Acad. Sci. U.S.A. 99:16899-16903(2002).

1061; PubMed=10085077;

hang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;

is angiogenesis and proliferation of endothelial cells.";

1. 274:8455-8459(1999).

; Binds to FN14 and possibly also to TNFRSF12/APO3. Weak

of apoptosis in some cell types. Mediates NF-kappaB

on. May promote angiogenesis and the proliferation of

al cells.

Homotrimer (Potential).

AR LOCATION: Type II membrane protein and secreted.

PECIFICITY: Highly expressed in adult heart, pancreas,

muscle, brain, colon, small intestine, lung, ovary,

spleen, lymph node, appendix and peripheral blood

es. Low expression in kidney, testis, liver, placenta,

id bone marrow. Also detected in fetal kidney, liver,

brain.

CC -!- PTM: The soluble form derives from the membrane form

CC by proteolytic processing.

CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.

CC -!- CAUTION: Ref.3 sequence differs from that shown due to a

CC frameshift in position 125.

CC

CC This SWISS-PROT entry is copyright. It is produced through a coll

CC between the Swiss Institute of Bioinformatics and the EMBL ou

CC the European Bioinformatics Institute. There are no restriction

CC use by non-profit institutions as long as its content is n

CC modified and this statement is not removed. Usage by and for

CC entities requires a license agreement (See <http://www.isb-sib.ch>,

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; AF030099; AAC51923.1; -

CC EMBL; AF055872; AAC39724.1; -

CC EMBL; BC019047; AAH19047.1; AUT_FRAME.

CC Genew; HGNC:11927; TNFSF12.

CC MIM; 602695; -

CC

CC GO; GO:0005887; C:integral to plasma membrane; TAS.

CC GO; GO:0005102; F:receptor binding; TAS.

CC GO; GO:0006917; P:induction of apoptosis; TAS.

CC GO; GO:0007165; P:signal transduction; TAS.

CC InterPro; IPR006052; TNF family.

CC InterPro; IPR008983; TNF_like.

CC Pfam; PF00229; TNF; 1.

CC SMART; SM00207; TNF; 1.

CC PROSITE; PS00251; TNF_1; FALSE_NEG.

CC PROSITE; PS0049; TNF_2; 1.

CC Cytokine; Angiogenesis; Apoptosis; Transmembrane; Glycoprotein;

KW signal-anchor.

KW CHAIN 1 249 TUMOR NECROSIS FACTOR LIGAND SUPERF

FT CHAIN 94 249 MEMBER 12, MEMBRANE FORM

FT CHAIN 94 249 TUMOR NECROSIS FACTOR LIGAND SUPERF

FT DOMAIN 1 21 MEMBER 12, SECRETED FORM.

FT TRANSMEM 22 42 CYTOPLASMIC (POTENTIAL).

FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO

FT SIGNAL-ANCHOR (POTENTIAL).

FT DOMAIN 43 249 EXTRACELLULAR (POTENTIAL).

FT SITE 93 94 CLEAVAGE.

FT CARBOHYD 139 139 N-LINKED (GLCNAC...).

FT SEQUENCE 249 AA; 27216 MW; E660843361C28EBA CRC64;

SQ

Query Match 87.7%; Score 249; DB 1; Length 249;

Best Local Similarity 100.0%; Pred. No. 2.9e-232;

Matches 249; Conservative 0; Mismatches 0; Indels 0;

QY 36 MAARRSQRRRGRGEPGTALLVPLALGIGLALACIGLLAVVSLGSRASLSAQEP

Db 1 MAARRSQRRRGRGEPGTALLVPLALGIGLALACIGLLAVVSLGSRASLSAQEP

QY 96 VAEDDDSEINLPQTEESQDPAPFLNLRPRRSPAPKGRKTRARRAIAAHVEVHP

Db 61 VAEDDDSEINLPQTEESQDPAPFLNLRPRRSPAPKGRKTRARRAIAAHVEVHP

QY 156 GAQAGVDGTSGWEERINSSPLRYNRQIGEFIVTRAGLYLYCQVHFDEGKAV

Db 121 GAQAGVDGTSGWEERINSSPLRYNRQIGEFIVTRAGLYLYCQVHFDEGKAV

QY 216 LLVDGVIALRCLEFSAATAASLGPQLRCQVSGLLALRPOSSLRIRTPWAHLK

Db 181 LLVDGVIALRCLEFSAATAASLGPQLRCQVSGLLALRPOSSLRIRTPWAHLK

QY 276 TYFGLFQVH 284

Db 241 TYFGLFQVH 249

RESULT 2

TN12_MOUSE

ID TN12_MOUSE STANDARD; PRT; 225 AA.

AC OS4907; Q9CTP2;

DT 28-FEB-2003 (Rel. 41, Created)

3.5%; Score 10; DB 1; Length 111;
 100.0%; Pred. No. 0.038; 0; Indels 0; Gaps 0;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALACGL 72
 |||||
 ALACGL 49

STANDARD; PRT; 733 AA.

(Rel. 33, Created)
 (Rel. 33, Last sequence update)
 (Rel. 42, Last annotation update)
 I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B).
 sensis (Marine centric diatom).

tramenophiles; Bacillariophyta; Coscinodiscophyceae;
 ycidae; Eupodiscaceae; Eupodiscaceae; Odontella.
 339;

N.A.
 .. Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
 last genome of a chlorophyll a+c-containing alga,
 sensis"; J. Phycol. 13:336-342(1995).

: PsaA and PsaB bind P700, the primary electron donor of
 tem I (PSI), as well as the electron acceptors A0, A1 and
 is a plastocyanin/cytochrome c6-ferredoxin oxidoreductase,
 ng photonic excitation into a charge separation, which
 s an electron from the donor P700 chlorophyll pair to the
 copically characterized acceptors A0, A1, FX, FA and FB in
 ized P700 is reduced on the luminal side of the
 d membrane by plastocyanin or cytochrome c6.

: P700 is a chlorophyll A/chlorophyll A' dimer, A0 is one
 chlorophyll A, A1 is one or both phylloquinones and FX is
 4Fe-4S iron-sulfur center (By similarity).
 The psaa/B heterodimer binds the P700 chlorophyll special
 subsequent electron acceptors. PSI consists of a core
 complex that captures photons, and an electron transfer
 at converts photonic excitation into a charge separation.
 ytic PSI reaction center is composed of at least 11
 (By similarity).

LAR LOCATION: Integral membrane protein. Chloroplast
 d membrane (By similarity).
 TY: Belongs to the psaa/psaB family.

ROT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 -profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 uires a license agreement (see <http://www.isb-sib.ch/announce/>
 mail to license@isb-sib.ch).

; CAA91749.1; -.

S78376.

; IJBO.

482; -; 1.

R006244; PsaB.

R001280; PSI_PsaA/B.

3; psaa psaa; 1.

257; PHOTOSYSAAB.

GR01336; PsaB; 1.

0419; PHOTOSYSTEM I_PSAAB; 1.

electron transport; Photosynthesis; Thylakoid;

I; Chlorophyll; Metal-binding; Iron; Magnesium;

4Fe-4S; Transmembrane; Chloroplast.

46

I (POTENTIAL).

FT TRANSMEM 134 157
 FT TRANSMEM 174 198
 FT TRANSMEM 272 230
 FT TRANSMEM 329 352
 FT TRANSMEM 368 394
 FT TRANSMEM 416 438
 FT TRANSMEM 516 534
 FT TRANSMEM 574 595
 FT TRANSMEM 642 664
 FT TRANSMEM 706 726
 FT METAL 558
 FT METAL 567 567
 FT METAL 653 653
 FT METAL 661 661
 FT BINDING 669 669
 FT BINDING 670 670
 FT SEQUENCE 733 AA; 82103 MW; 13439AF1E441BBF7 CRC64;
 SQ

Query Match 3.2%; Score 9; DB 1; Length 733;

Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 9; Conservative 0; Mismatches 0; Indels 0;

QY 63 LGLALACLG 71

Db 333 LGLALACLG 341

RESULT 5

PSAB CVACA

ID PSAB CVACA STANDARD; PRT; 734 AA.

AC Q9TLQ6;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Photosystem I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B).

GN PSAB

OS Cyanidium caldarium.

OG Chloroplast.

OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridi

OC Cyanidium.

OX NCBI_TaxID=2771;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RK-1;

RX MEDLINE=20496959; PubMed=11040290;

RA Gloeckner G., Rosenthal A., Valentin K.-U.;

RT "The structure and gene repertoire of an ancient red algal plast

genome.";

RL J. Mol. Evol. 51:382-390(2000).

CC -!- FUNCTION: PsaA and PsaB bind P700, the primary electron dono

CC Photosystem I (PSI), as well as the electron acceptors A0, A

CC FX. PSI is a plastocyanin/cytochrome c6-ferredoxin oxidoredu

CC converting photonic excitation into a charge separation, whi

CC transfers an electron from the donor P700 chlorophyll pair t

CC spectroscopically characterized acceptors A0, A1, FX, FA and

CC turn. Oxidized P700 is reduced on the luminal side of the

CC thylakoid membrane by plastocyanin or cytochrome c6.

CC -!- COFACTOR: P700 is a chlorophyll A/chlorophyll A' dimer, A0 i

CC or more chlorophyll A, A1 is one or both phylloquinones and

CC a shared 4Fe-4S iron-sulfur center (By similarity).

CC -!- SUBUNIT: The psaa/B heterodimer binds the P700 chlorophyll s

CC pair and subsequent electron acceptors. PSI consists of a co

CC antenna complex that captures photons, and an electron trans

CC chain that converts photonic excitation into a charge separa

CC The eukaryotic PSI reaction center is composed of at least 1

CC subunits (By similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast

CC Thylakoid membrane (By similarity).

CC -!- SIMILARITY: Belongs to the psaa/psaB family.

entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements>).

AAFL12881.1; -
LJB0.
120.-; 1.
06244; PsaB.
01280; PSI_PsaA/B.
psaA_psaB; 1.
7; PHOTOSYSAAB.
01336; psaB; 1.
19; PHOTOSYSTEM I PSAB; 1.
Iron transport; Photosynthesis; Thylakoid;
Chlorophyll; Metal-binding; Iron; Magnesium;
Fe-4S; Transmembrane; Chloroplast.
6 69 I (POTENTIAL).
15 158 II (POTENTIAL).
75 199 III (POTENTIAL).
73 291 IV (POTENTIAL).
10 353 V (POTENTIAL).
19 395 VI (POTENTIAL).
7 439 VII (POTENTIAL).
7 535 VIII (POTENTIAL).
75 596 IX (POTENTIAL).
13 665 X (POTENTIAL).
7 727 XI (POTENTIAL).
19 559 IRON-SULFUR (4FE-4S) (SHARED WITH DIMERIC PARTNER) (BY SIMILARITY).
18 568 IRON-SULFUR (4FE-4S) (SHARED WITH DIMERIC PARTNER) (BY SIMILARITY).
14 654 MAGNESIUM (CHLOROPHYLL-A B1 AXIAL LIGAND; P700 SPECIAL PAIR) (BY SIMILARITY).
12 662 MAGNESIUM (CHLOROPHYLL-A B3 AXIAL LIGAND) (BY SIMILARITY).
10 670 CHLOROPHYLL-A B3 (BY SIMILARITY).
1 671 PHYLOQUINONE B (BY SIMILARITY).
AA; 82359 MW; 4496AA2AE59CA9B9 CRC64;

3.2%; Score 9; DB 1; Length 734;
identity 100.0%; Pred. No. 1.8;
conservative 0; Mismatches 0; Indels 0; Gaps 0;
ACLG 71
ACLG 342

STANDARD; PRT; 179 AA.

rel. 35, Created)
rel. 35, Last sequence update)
rel. 41, Last annotation update)
ogenase 15 kDa subunit precursor (G3-ADH subunit III).
oxydase (Gluconobacter suboxydase).
eobacteria; Alphaproteobacteria; Rhodospirillales;
ae; Gluconobacter.

N.A., AND SEQUENCE OF 26-40.

328;
225; PubMed=9055427;
inouchi S.;
ion of the genes encoding the three-component membrane-dehydrogenase from Gluconobacter suboxydase and their

expression in Acetobacter pasteurianus.";
Appl. Environ. Microbiol. 63:1131-1138(1997).
-1- FUNCTION: NOT ESSENTIAL FOR ALCOHOL DEHYDROGENASE ACTIVITY.
CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Membrane-bound, facing the periplasmic
(Potential).

This SWISS-PROT entry is copyright. It is produced through a coll.
between the Swiss Institute of Bioinformatics and the EMBL out
the European Bioinformatics Institute. There are no restriction
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for c
entities requires a license agreement (See <http://www.isb-sib.ch/>
or send an email to license@isb-sib.ch).

EMBL; D86440; BAA19756.1; -
Membrane; Periplasmic; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 179 ALCOHOL DEHYDROGENASE 15 kDa SUBUNIT
FT MOD RES 25 25 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 179 AA; 19943 MW; F6AF243656B3CC66 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G.

QY 59 LALGLGLA 66
Db 11 LALGLGLA 18
|||||

RESULT 7

Y304 BRUME
ID Y304 BRUME STANDARD; PRT; 220 AA.
AC CAYD73;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein BMEI10304.
GN BMEI10304.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kaprat V., Redkar R.J., Patra G., Mujter C., Lo
Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzma
Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
Haselkorn R., Kyriades N., Overbeek R.;
RT The genome sequence of the facultative intracellular pathogen
Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the UPF0191 family.

This SWISS-PROT entry is copyright. It is produced through a coll.
between the Swiss Institute of Bioinformatics and the EMBL out
the European Bioinformatics Institute. There are no restriction
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for c
entities requires a license agreement (See <http://www.isb-sib.ch/>
or send an email to license@isb-sib.ch).

EMBL; AE009669; AAL53546.1; -
PIR; AG3547; AG3547.
HAMAP; MF_01207; -; 1.
InterPro; IPR007916; UPF0191.
Pfam; PF05252; UPF0191; 1.

protein; Transmembrane; Complete proteome.

20 39 POTENTIAL.
54 72 POTENTIAL.
85 104 POTENTIAL.
124 146 POTENTIAL.
153 175 POTENTIAL.
179 198 POTENTIAL.
20 AA; 24815 MW; 182C0244743B17FA CRC64;

2.8%; Score 8; DB 1; Length 220;
arity 100.0%; Pred. No. 5.8;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/PLAL 61

/PLAL 139

STANDARD; PRT; 220 AA.

Rel. 43, Created)

Rel. 43, Last sequence update)

Rel. 43, Last annotation update)

protein BRA0991.

Neobacteria; Alphaproteobacteria; Rhizobiales;

Brucella.

1461;

1 N.A.

1 Biovar 1;

741; PubMed-12271122;

Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,

Johnson R.J., Umayam L., Brinkac L.M., Beanan M.J.,

Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,

Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,

Tettelin H., Gill S.R., White O., Salzberg S.L.,

Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.,

antis genome reveals fundamental similarities between

ant pathogens and symbionts."

Acad. Sci. U.S.A. 99:13148-13153(2002).

AR LOCATION: Integral membrane protein (Potential).

Y: Belongs to the UPF0191 family.

NOT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
dires a license agreement (See <http://www.isb-sib.ch/announce/>
mail to license@isb-sib.ch).

12; AAN34160.1; -.

107; -; 1.

1007916; UPF0191.

1; UPF0191; 1.

protein; Transmembrane; Complete proteome.

20 39 Potential.

54 72 Potential.

85 104 Potential.

124 146 Potential.

153 175 Potential.

179 198 Potential.

20 AA; 24796 MW; AC2C060433169497 CRC64;

2.8%; Score 8; DB 1; Length 220;

arity 100.0%; Pred. No. 5.8;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 ALLVPLAL 61
DB 132 ALLVPLAL 139
RESULT 9
MSHR PANTR
ID MSHR PANTR STANDARD; PRT; 317 AA.
AC Q9TUK4; O864L1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Melanocyte stimulating hormone receptor (MSH-R) (Melanotropin
DE receptor) (Melanocortin-1 receptor) (MCI-R).
GN MCI-R.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Rees J.L., Harding R.M., Healy E., Jackson I.J., Ray A.J., Ellis
RA Flanagan N., Todd C., Dixon C., Matthews J.N., Sajantila A.,
RA Birch-Machin M.A.;
RT "Chimpanzee melanocortin 1 sequence."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate 3;
RX MEDLINE=22572539; PubMed=12687585;
RA Mundy N.I., Kelly J.;
RT "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
RT primates."
RL Am. J. Phys. Anthropol. 121:67-80(2003).
CC -!- FUNCTION: Receptor for MSH (alpha, beta and gamma) and ACTH.
CC activity of this receptor is mediated by G proteins which act
CC adenylylate cyclase (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled recepto.
CC This SWISS-PROT entry is copyright. It is produced through a col-
CC between the Swiss Institute of Bioinformatics and the EMBL out-
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for
CC entites requires a license agreement (See <http://www.isb-sib.ch>,
CC or send an email to license@isb-sib.ch).

EMBL; AJ245705; CAB53398.1; -.
EMBL; AY205086; AAP30960.1; -.
InterPro; IPR000276; GPCR_Rhodpsn.
Pfam; PF00001; 7tm.1; 1.
PRINTS; PR00237; GPCRHOODPSN
PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
PROSITE; PS00262; G-PROTEIN RECEPTOR_F1_2; 1.
G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 38 63 1 (POTENTIAL).
FT DOMAIN 64 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 93 2 (POTENTIAL).
FT DOMAIN 94 118 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 119 140 3 (POTENTIAL).
FT DOMAIN 141 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 183 4 (POTENTIAL).
FT DOMAIN 184 191 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 192 211 5 (POTENTIAL).
FT DOMAIN 212 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 266 6 (POTENTIAL).
FT DOMAIN 267 279 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 280 300 7 (POTENTIAL).
FT DOMAIN 301 317 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 N-LINKED (GLCNAC...) (POTENTIAL).

5 315 S-palmitoyl cysteine (Potential).
AA; 34699 MW; 5615D2146E1D247F CRC64;
rity 100.0%; Score 8; DB 1; Length 317;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
IAA 144
IAA 167

STANDARD; PRT; 379 AA.

rel. 35, Created)
rel. 35, Last sequence update)
rel. 43, Last annotation update)
the dehydrogenase beta chain (EC 1.2.1.2).
jannaschii.
rchaeta; Methanococci; Methanococcales;
ccaceae; Methanocaldococcus.
0;
N.A.
DSM 2661 / ATCC 43067;
99; PubMed=6688087;
te O.; Olsen G.J.; Zhou L.; Fleischmann R.D.,
lake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
irkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
oghagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
oberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
raser C.M., Smith H.O., Woese C.R., Venter J.C.;
me sequence of the methanogenic archaeon, Methanococcus
58-1073(1996).
ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
Binds 2 4Fe-4S clusters (Probable).
imer of alpha and beta chains (By similarity).
: The iron-sulfur centers are similar to those of
type 4Fe-4S ferredoxins.
: STRONG, TO THE BETA SUBUNIT OF M.THERMOAUTOTROPHICUM

T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See <http://www.isb-sib.ch/announce/>
il to license@isb-sib.ch).

AAB37986.1; -
64300.

01450; 4Fe4S ferredoxin.
07525; FrhB_FdhB_C.
07516; FrhB_FdhB_N.
fer4; 1.
FrhB_FdhB_C; 1.
FrhB_FdhB_N; 1.
98; 4FE4S_FERREDOXIN; 2.
rotein; Oxidoreductase; NAD; Electron transport;
Fe-4S; Complete proteome.
0 280 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
3 283 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
6 286 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
0 290 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
0 330 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 333 333 IRON-SULFUR (4FE-4S) (BY SIMILARITY)
FT METAL 336 336 IRON-SULFUR (4FE-4S) (BY SIMILARITY)
FT METAL 340 340 IRON-SULFUR (4FE-4S) (BY SIMILARITY)
SQ SEQUENCE 379 AA; 43014 MW; 9C257CCAD5547F5A CRC64;
Query Match 2.8%; Score 8; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 217 LVDGVLLAL 224
Db 35 LVDGVLLAL 42

RESULT 11
ZP3_MESAU STANDARD; PRT; 422 AA.
ID_ZP3_MESAU
AC P23491;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Zona pellucida sperm-binding protein 3 precursor (Zona pellucida
glycoprotein ZP3) (Sperm receptor) (Zona pellucida protein C).
GN ZP3.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Ovary;
RC MEDLINE=91078540; PubMed=2257975;
RA Kinloch R.A., Ruiz-Seller B., Wassarman P.M.;
RT "Genomic organization and polypeptide primary structure of zona
pellucida glycoprotein HZP3, the hamster sperm receptor.";
RL Dev. Biol. 142:414-421(1990).
CC -!- FUNCTION: Functions as a sperm-receptor. It is responsible for
sperm-adhesion to the zona pellucida, and may contribute to ti
species-specificity of the insemination.
CC -!- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1
matrix.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
matrix.
CC -!- TISSUE SPECIFICITY: Oocytes.
CC -!- DEVELOPMENTAL STAGE: GROWING OOCYTES.
CC -!- PTM: Sulfated glycoprotein with O-linked oligosaccharides.
CC -!- SIMILARITY: Contains 1 ZP domain.

This SWISS-PROT entry is copyright. It is produced through a colla
between the Swiss Institute of Bioinformatics and the EMBL out
the European Bioinformatics Institute. There are no restrictions
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for co
entries requires a license agreement (See <http://www.isb-sib.ch/>
or send an email to license@isb-sib.ch).

EMBL; M63629; AAA37079.1; -
InterPro; IPR001507; Endoglin/CD105.
Pfam; PF00100; zona_pellucida; 1.
PRINTS; PRO0023; ZPELLUCIDA.
SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
Extracellular matrix.
FT SIGNAL 1 22
FT CHAIN 23 422
FT DOMAIN 23 386
FT TRANSMEM 387 407
FT DOMAIN 408 422
FT DOMAIN 45 306
FT DOMAIN 119 158
FT DOMAIN 208 257
POTENTIAL.
ZONA PELLUCIDA SPERM-BINDING PROTEIN
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
ZP.
PRO-RICH.
PRO-RICH.

146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 22 AA; 45827 MW, D0F95B7FFBE7E01 CRC64;

2.8%; Score 8; DB 1; Length 422;
 larity 100.0%; Pred. No. 10;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;

ILGIA 66

|||||

ILGIA 393

STANDARD; PRT; 576 AA.

(Rel. 32, Created)
 (Rel. 32, Last sequence update)
 (Rel. 41, Last annotation update)
 p-binding protein cydC.

36.
 influenzae;
 teobacteria; Gammaproteobacteria; Pasteurellales;
 ae; Haemophilus.

4 N.A.

KW20 / ATCC 51907;

3630; PubMed:7542800;

3.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 3., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 , Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 3., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 itchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

a random sequencing and assembly of Haemophilus influenzae

196-512(1995).

: SOMEHOW INVOLVED IN THE CYTOCHROME D BRANCH OF AEROBIC
 ION. SEEMS TO BE A COMPONENT OF A TRANSPORT SYSTEM

LARITY).

AR LOCATION: Integral membrane protein. Inner membrane

3). Belongs to the ABC transporter family. MsbA subfamily.

ROT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 -profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 izes a license agreement (See <http://www.isb-sib.ch/announce/>
 mail to licenses@isb-sib.ch).

3 AAC22811.1; -

E64186.

3003593; AAA ATPase.

3001140; ABC_TM transp.

3003439; ABC transporter.

4; ABC membrane; 1.

5; ABC_tran; 1.

3006; ABC transporter; 1.

32; AAA; 1.

3929; ABC_TMIF; 1.

3211; ABC_TRANSPORTER_1; 1.

3893; ABC_TRANSPORTER_2; 1.

Transport; transmembrane; Inner membrane;

ecome.

FT TRANSMEM 16 36 POTENTIAL.
 FT TRANSMEM 38 58 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 244 264 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 FT DOMAIN 338 374 ABC TRANSPORTER.
 FT NP_BIND 372 379 ATP (POTENTIAL).
 SQ SEQUENCE 576 AA; 64831 MW; A9ACD8B9B294B1B3 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 576;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 58 PLALGLGL 65

|||||

Db 159 PLALGLGL 166

RESULT 13

GGT5_HUMAN

ID GGT5_HUMAN STANDARD; PRT; 586 AA.

AC P36269; Q96FC1; Q9UFM5;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Gamma-glutamyltransferase 5 precursor (EC 2.3.2.2) (Gamma-glutamyltransferase 5) (Gamma-glutamyltransferase-like activi-
 (GGT-rel).

DE (GGT-rel).

GN GGT1A1 OR GGT5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

SEQUENCE FROM N.A. (ISOFORM 1).

TISSUE=Placenta;

RX MEDLINE=91296809; PubMed=1676842;

RA Heisterkamp N., Rajpert-De Meyts E., Uribe L., Forman H.J.,

RA Groffen J.,

RT "Identification of a human gamma-glutamyl cleaving enzyme relate
 but distinct from, gamma-glutamyl transpeptidase."

RL Proc. Natl. Acad. Sci. U.S.A. 88:6303-6307(1991).

RN [2]

SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Petal kidney;

RA Blum H., Bauersachs S., Mewes H.-W., Gassenhuber J., Wiemann S.;

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

SEQUENCE FROM N.A.

RX MEDLINE=20057165; PubMed=10591208;

RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,

RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,

RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,

RA Bird C.P., Blakey S.B., Bridgman A.M., Buck D., Burgess J.,

RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark I

RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,

RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson J

RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A

RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,

RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall

RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,

RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,

RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,

RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,

RA McClay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.

RA Odeil C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C

RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross

RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,

RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.

RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,

RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming I

RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimi:

Kawasaki K., Sasaki T., Asakawa S., Kudoh J., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S., Chu L., Crabtree J., Deschamps S., Do A., Do T., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L., Wang Z., White J., Willingham D., Wu H., Yao Z., Y., Chisoe S., Murray J., Miller N., Mink P., nson D., Bemis G., Bentley D., Bradshaw H., Bourne S., Z., Fulton L., Goela D., Graves T., Hawkins J., K., Latreille P., Layman D., Ozersky P., Rohlfing T., lker C., Wamley A., Wohlmann P., Pepin K., Nelson J., ll J.A., Hillier L.W., Mardis E., Waterston R., ancel B.S., Shaikh T., Kurahashi H., Saitta S., uermid H.E., Johnson A., Wong A.C.C., Morrow B.E., kim U.J., Shizuya H., Simon M.I., Dumanski J.P., edra D., Seroussi E., Franconi I., Tapia I., Bruder C.E., Wilkenson P., Bodenteich A., Hartman K., Hu X., ie L., Tiliahun Y., Wright H.;
ence of human chromosome 22.";
3-495(1999).

N.A. (ISOFORM 1).

257; PubMed=12477932;
J., Feingold E.A., Grouse L.H., Derge J.G., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Jordan H., Moore T., Max S.I., Wang J., Heieh F., Marusina K., Farmer A.A., Rubin G.M., Hong L., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E., Udén T.B., Toshiyuki S., Carninci P., Prange C., Ruellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Ewan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Con E., Kettman M., Madan A., Rodrigues S., Sanchez A., adan A., Young A.C., Shevchenko Y., Bouffard G.G., Touchman J.W., Green E.D., Dickson M.C., Grimwood J., Schmutz J., Myers R.M., S.N., Krzywinski M.I., Skalska U., Smalish D.E., Schein J.E., Jones S.J.M., Marra M.A.;
ad initial analysis of more than 15,000 full-length
ed cDNA sequences.";
ad. Sci. U.S.A. 99:16899-16903(2002).
cleaves the gamma-glutamyl peptide bond of glutathione
s, but maybe not glutathione itself. Converts
le C4 (LTC4) to leukotriene D4 (LTD4).
ACTIVITY: (S-L-glutamyl)-peptide + an amino acid =
S-L-glutamyl-amino acid.
glutathione metabolism.
leukotriene metabolism; second step.
eterodimer composed of the light and heavy chains.
e site is located in the light chain (By similarity).
AR LOCARION: Type II membrane protein (By similarity).
VE PRODUCTS:
ernative splicing; Named isoforms=2;
36269-1; Sequence=Displayed;
36269-2; Sequence=VSP_008146;
experimental confirmation available;
y: Belongs to the gamma-glutamyltransferase family.
Ref.2 sequence differs from that shown due to a
t in position 446.

DT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See <http://www.isb-sib.ch/announce/>)
ail to license@isb-sib.ch).

CC EMBL; M64099; AAA58503.1; --
DR EMBL; AL117414; CAB55910.1; ALT FRAME.
DR EMBL; AP000354; -- NOT ANNOTATED_CDS.
DR EMBL; BC011362; AAH11362.1; --
DR PIR; A41125; A41125.
DR PIR; T17220; T17220.
DR MEROPS; T03.002; --
DR Genew; HGNC:4260; GGTAL1.
DR MIM; 137168; --
DR GO; GO:0016021; C: integral to membrane; TAS.
DR GO; GO:0003840; F: gamma-glutamyltransferase activity; TAS.
DR GO; GO:0006520; P: amino acid metabolism; TAS.
DR GO; GO:0006749; P: glutathione metabolism; TAS.
DR InterPro; IPR000101; Peptidase T3.
DR Pfam; PF01019; G: glutathione transferase; 1.
DR PRINTS; PS01210; GGTTRANSFERASE.
DR PROSITE; PS00462; G: GLO-TRANSFERASE; 1.
KW Glutathione biosynthesis; Leukotriene biosynthesis; Transferase;
KW Acyltransferase; Signal-anchor; Transmembrane; Zymogen; Glycoprot
KW Alternative splicing.
FT CHAIN 1 387 GAMMA-GLUTAMYLTRANSFERASE 5 HEAVY
FT CHAIN 388 586 CHAIN (BY SIMILARITY).
FT DOMAIN 1 8 GAMMA-GLUTAMYLTRANSFERASE 5 LIGHT
FT TRANSMEM 9 29 CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROT
(POTENTIAL).
FT DOMAIN 30 586 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 535 535 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 101 132 Missing (in isoform 2).
FT CONFLICT 330 330 R -> K (IN REF. 2 AND 4).
FT CONFLICT 408 408 N -> Y (IN REF. 2).
FT CONFLICT 437 437 W -> R (IN REF. 2 AND 4).
SQ SEQUENCE 586 AA; 62319 MW; 1BE543CB0934B16B CRC64;
Query Match 2.8%; Score 8; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G
QY 61 LGLGLALA 68
Db 14 LGLGLALA 21
RESULT 14
HSP3 OCTVU
ID_HSP3 OCTVU STANDARD; PRT; 24 AA.
AC P83215;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sperm protamine P3 (Po3) (Fragment).
OS Octopus vulgaris (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoide
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=6645;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=Sperm;
RA Gimenez-Bonafe P., Ribes E., Buesa C., Sautiere P., Kouach M.,
RA Ausio J., Kasinsky H.E., Chiva M.;
RT "Chromatin remodelling and protamines during spermiogenesis of Oc
RT vulgaris (Cephalopoda)".
RL J. Exp. Zool. 0:0-0(2001).
CC -!- FUNCTION: Protamines substitute for histones in the chromatin
sperm during the haploid phase of spermatogenesis. They compa

A into a highly condensed, stable and inactive

LAR LOCATION: Nuclear.

PECIFICITY: Testis.

CTROMETRY: MW=4389; METHOD=Electrospray.

36; C:nucleosome; NAS.

34; C:nucleus; NAS.

77; F:DNA binding; NAS.

01; P:chromosome organization and biogenesis (sen. . .; NAS.

76; P:mitotic chromosome condensation; NAS.

34; P:nucleosome assembly; NAS.

33; P:spermatogenesis; NAS.

protein; Nucleosome core; Spermatogenesis;

DNA condensation; Nuclear protein.

1 16 POLY-ARG.

24 24

4 AA; 3381 MW; 308E90ED9D2C9C9C CRC64;

2.5%; Score 7; DB 1; Length 24;

larity 100.0%; Pred. No. 7.7;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3RRG 49

||||

3RRG 17

STANDARD; PRT; 30 AA.

(Rel. 41, Created)

(Rel. 41, Last sequence update)

(Rel. 41, Last annotation update)

ine P5 (Po5).

aris (Octopus).

stazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;

es; Octopoda; Incirrata; Octopodidae; Octopus.

645;

NCTION, AND MASS SPECTROMETRY.

fe P., Ribes E., Buesa C., Sautiere P., Kouach M.,

sinsky H.E., Chiva M.;

emodelling and protamines during spermiogenesis of Octopus

phalopoda).";

0.0-0.0(2001).

: Protamines substitute for histones in the chromatin of

ring the haploid phase of spermatogenesis. They compact

A into a highly condensed, stable and inactive

LAR LOCATION: Nuclear.

PECIFICITY: Testis.

CTROMETRY: MW=3941; METHOD=Electrospray.

36; C:nucleosome; NAS.

34; C:nucleus; NAS.

77; F:DNA binding; NAS.

01; P:chromosome organization and biogenesis (sen. . .; NAS.

76; P:mitotic chromosome condensation; NAS.

34; P:nucleosome assembly; NAS.

33; P:spermatogenesis; NAS.

protein; Nucleosome core; Spermatogenesis;

DNA condensation; Nuclear protein.

2 15 POLY-ARG.

17 26

0 AA; 3943 MW; 14F1BC7E4D277049 CRC64;

2.5%; Score 7; DB 1; Length 30;

larity 100.0%; Pred. No. 9.3;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3RRG 49

||||

Db 5 RRRGRRG 11

RESULT 16

HSPI_SAGIM

ID HSPI_SAGIM STANDARD; PRT; 49 AA.

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Sperm protamine P1 (Cysteine-rich protamine).

GN PRM1.

OS Saginus imperator (Emperor tamarin).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi

OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Sagu

OX NCBI_TaxID=9491;

RN [1]

SEQUENCE FROM N.A.

RP TISSUE=Liver;

RX MEDLINE=92051332; PubMed=1840669;

RA Queralt R., Oliva R.;

RT "Protamine 1 gene sequence from the primate Saginus imperator

isolated with PCR using consensus oligonucleotides.";

RL Nucleic Acids Res. 19:5786-5786(1991).

CC -!- FUNCTION: Protamines substitute for histones in the chromati

sperm DNA into a highly condensed, stable and inactive compl

-!- SUBUNIT: Cross-linked by interchain disulfide bonds around t

DNA-helix (By similarity).

-!- SUBCELLULAR LOCATION: Nuclear.

-!- TISSUE SPECIFICITY: Testis.

-!- SIMILARITY: Belongs to the protamine P1 family.

CC This SWISS-PROT entry is copyright. It is produced through a col

between the Swiss Institute of Bioinformatics and the EMBL ou

the European Bioinformatics Institute. There are no restrictio

use by non-profit institutions as long as its content is i

modified and this statement is not removed. Usage by and for

entities requires a license agreement (See <http://www.isb-sib.ch>

or send an email to license@sib-sib.ch).

CC -----

EMBL; X61678; CAA43853.1; --

DR PIR; S22582; S22582.

DR InterPro; IPR000221; Protamine_P1.

DR Pfam; PF00260; Protamine_P1; 1.

DR PROSITE; PS00048; PROTAMINE_P1; 1.

KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-bind

KW Testis; DNA condensation; Nuclear protein.

FT INIT_MET 0 0

SQ SEQUENCE 49 AA; 6545 MW; 8389C403F5B207F6 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 49;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 42 QRRRGR 48

|||||

Db 17 QRRRGR 23

RESULT 17

HSPI_DIDMA

ID HSPI_DIDMA STANDARD; PRT; 57 AA.

AC P35305;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Sperm protamine P1.

GN PRM1.

OS Didelphis marsupialis virginiana (North American opossum), and

OC Monodelphis domestica (Short-tailed grey opossum).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi

OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.

57, 13616;
 N.A.
 supialis;
 300; PubMed=9344286;
 Nishikawa S., Connor W., Dixon G.H.;
 tion of a marsupial sperm protamine gene and its
 om the North American opossum (Didelphis
 ';
 m. 215:63-72 (1993).
 N.A.
 istica;
 51; PubMed=7700877;
 Grajewski C., Western M., Winkfein R.J., Dixon G.H.;
 logeny and evolution of marsupial protamine P1 genes.;
 Lond., B. Biol. Sci. 259:7-14 (1995).
 Protamines substitute for histones in the chromatin of
 ing the haploid phase of spermatogenesis. They compact
 into a highly condensed, stable and inactive complex.
 AR LOCATION: Nuclear.
 SPECIFICITY: Testis.
 (: Belongs to the protamine P1 family.
 JT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 uses a license agreement (See <http://www.isb-sib.ch/announce/>
 ail to license@isb-sib.ch).

 AAA02812.1; -;
 CAA52193.1; -;
 AAA74612.1; -;
 i34045.
 00221; Protamine P1.
 protamine P1; 1.
 148; PROTAMINE P1; 1.
 corein; Nucleosome core; Spermatogenesis; DNA-binding;
 ndensation; Nuclear protein.
 0
 0 BY SIMILARITY.
 AA; 7810 MW; 283715280214E52 CRC64;

 2.5%; Score 7; DB 1; Length 57;
 urity 100.0%; Pred No. 16;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 RG 49
 ||
 RG 40

 STANDARD; PRT; 115 AA.
 Rel. 40, Created)
 Rel. 40, Last sequence update)
 Rel. 42, Last annotation update)
 id protein Acp62F precursor.
 162.
 lanogaster (Fruit fly).
 :azoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 pterygota; Diptera; Brachycera; Muscomorpha;
 prosophilidae; Drosophila.
 ?;
 N.A.; FUNCTION, AND TISSUE SPECIFICITY.
 -S; TISSUE=Male accessory gland;
 20; PubMed=9474779;
 Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,
 ig Y.O., Neubaum D.M., Park M., Tram U.K.;

RT
 RT melanogaster.";
 RL Insect Biochem. Mol. Biol. 27:825-834 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N
 RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L
 RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin
 RA Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dun
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman
 RA Fosler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 [3]
 RP SEQUENCE OF 7-111 FROM N.A.
 RC STRAIN=ZIM62H-12C, ZIM62H-16C, ZIM62H-28C, ZIM62H-30C, ZIM62H-34C
 RC ZIM62I-5C, ZIM62I-10C, ZIM62I-17C, ZIM62I-18C, and ZIM62I-53C;
 RX MEDLINE=20556153; PubMed=11102381;
 RA Begun D.J., Whitley P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;
 RT "Molecular population genetics of male accessory gland proteins i
 RT Drosophila.";
 RL Genetics 156:1879-1888 (2000).
 CC -I- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGE;
 CC MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL
 CC FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO
 CC AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STO
 CC AND EGG RELEASE.
 CC -I- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -I- TISSUE SPECIFICITY: Seminal fluid.
 CC -I- SIMILARITY: SOME, TO P.NIGRIVENTER TX2-6.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a colla
 CC between the Swiss Institute of Bioinformatics and the EMBL outst
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is in
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U85763; AAB96387.1; -;
 CC EMBL; AB003475; AAF47683.1; -;

```

8; AAG35367.1; -
9; AAG35368.1; -
0; AAG35369.1; -
1; AAG35370.1; -
2; AAG35371.1; -
3; AAG35372.1; -
4; AAG35373.1; -
5; AAG35374.1; -
6; AAG35375.1; -
7; AAG35376.1; -
10020509; ACP62F.
17; F:serine protease inhibitor activity; IDA.
0; P:determination of adult life span; NAS.
002919; TIL_Cyrich.
; TIL; 1.
nal.
1 24 POTENTIAL.
25 115 ACCESSORY GLAND PROTEIN ACP62F.
34 88 TIL.
5 AA; 12570 MW; 4326AA6FC32291D CRC64;
2.5%; Score 7; DB 1; Length 115;
arity 100.0%; Pred.No.30;
onservative 0; Mismatches 0; Indels 0; Gaps 0;
LLL 74
|||
LLL 17
STANDARD; PRT; 118 AA.
Rel. 01, Created)
Rel. 01, Last sequence update)
Rel. 41, Last annotation update)
ment ISS very hypothetical 12 kDa protein.
oli.
reobacteria; Gammaproteobacteria; Enterobacteriales;
aceae; Escherichia.
1 N.A.
1653; PubMed=6269959;
ahn M.;
de sequence of ISS from Escherichia coli.";
74(1981).
1 N.A.
1652; PubMed=6269958;
van Bree M.P.;
de sequence and protein-coding capability of the
element ISS.";
.63(1981).
1 N.A.
uncan M., Allen E., Araujo R., Aparicio A., Chung E.,
ierspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
Lew H., Lin D., Namath A., Oefner P., Roberts D.,
3P-1996) to the EMBL/GenBank/DBJ databases.
1 N.A.
Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
Inokuchi H., Maki T., Hatada E., Fukuda R.,
Mizuno T., Makino K., Nakata A., Yura T., Sampei G.,
3B-1996) to the EMBL/GenBank/DBJ databases.
1 N.A.

```

```

RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a col-
CC between the Swiss Institute of Bioinformatics and the EMBL out-
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is in-
CC modified and this statement is not removed. Usage by and for
CC entities requires a license agreement (See http://www.isb-sib.ch,
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; J01734; -; NOT ANNOTATED_CDS.
DR EMBL; U70214; AAB08680.1; -
DR EMBL; D83336; -; NOT ANNOTATED_CDS.
DR EMBL; D90771; BAA14925.1; -
DR EMBL; D90772; BAA14935.1; -
DR EMBL; D90831; BAA15715.1; -
DR EMBL; D90841; BAA15872.1; -
DR EMBL; D90847; BAA15958.1; -
DR EMBL; D90848; BAA15963.1; -
DR PIR; B91483; IEEC5B.
KW Hypothetical protein; Transposable element.
SQ SEQUENCE 118 AA; 12270 MW; 348014FAC765058E CRC64;
Query Match 2.5%; Score 7; DB 1; Length 118;
Best Local Similarity 100.0%; Pred.No.31;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 23 DGGAVRQ 29
|||
DB 107 DGGAVRQ 113
|||
RESULT 20
IL13 MOUSE
ID IL13 MOUSE STANDARD; PRT; 131 AA.
AC P20109.
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Interleukin-13 precursor (IL-13) (T-cell activation protein P600
GN IL13 OR IL-13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mu-
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89093958; PubMed=2521353;
RA Brown K.D., Zurawski S.M., Mosmann T.R., Zurawski G.;
RT "A family of small inducible proteins secreted by leukocytes are
RT members of a new superfamily that includes leukocyte and
RT fibroblast-derived inflammatory agents, growth factors, and
RT indicators of various activation processes.";
RL J. Immunol. 142:679-687(1989).
CC -!- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION
CC SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
CC MAY BE CRITICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPON-
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the IL-4 / IL-13 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a col-
CC between the Swiss Institute of Bioinformatics and the EMBL out-

```

bioinformatics Institute. There are no restrictions on its profit institutions as long as its content is in no way his statement is not removed. Usage by and for commercial res a license agreement (See <http://www.isb-sib.ch/announce/>).

 AAA40149.1; -.
 31TR.
 ; IL13.
 03634; Interleukin_13.
 01325; Interleukin_4_13.
 Interleukin_13; 1.
 ; IL4_13; 1.
 ; IL4_13; 1.
 138; INTERLEUKIN_4_13; 1.
 oprotein; Signal.
 1 21 BY SIMILARITY.
 2 131 INTERLEUKIN-13.
 3 179 BY SIMILARITY.
 4 93 BY SIMILARITY.
 5 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 6 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
 7 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
 AA; 14107 MW; 954F93F105713FED CRC64;

2.5%; Score 7; DB 1; Length 131;
 100.0%; Pred. No. 34;
 0; Mismatches 0; Indels 0; Gaps 0;

IG 71
 ||
 IG 15

STANDARD; PRT; 131 AA.

Rel. 32, Created)
 Rel. 32, Last sequence update)
 Rel. 36, Last annotation update)
 precursor (IL-13) (T-cell activation protein P600).

cus (Rat).
 azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 eria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 16;

N.A.
 ;Dawley; TISSUE=Kidney cortex;
 38; PubMed=7916615;

it interleukin-13 (IL-13) cDNA and analysis of IL-13
 in experimental glomerulonephritis";
 ws. Res. Commun. 197:612-618(1993).

CYTOKINE INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.
 WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
 TICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONSES
 ARITY).

AR LOCATION: Secreted.

); Belongs to the IL-4 / IL-13 family.

OT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation
 bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 1res a license agreement (See <http://www.isb-sib.ch/announce/>).

 AAA16478.1; -.
 152290.

DR HSP; P35225; 31TR.
 DR InterPro; IPR003634; Interleukin_13.
 DR InterPro; IPR001325; Interleukin_4_13.
 DR Pfam; PF03487; Interleukin_13; 1_
 DR ProDom; PD015987; Interleukin_13; 1.
 DR SMART; SM00190; IL4_13; 1.
 DR PROSITE; PS00838; INTERLEUKIN_4_13; 1.
 KW Cytokine; Glycoprotein; Signal.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT CHAIN 22 131 INTERLEUKIN-13.
 FT DISULFID 52 80 BY SIMILARITY.
 FT DISULFID 68 94 BY SIMILARITY.
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 131 AA; 14093 MW; E5008CAB8DE8C201 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LALACLG 71
 |||||
 DB 9 LALACLG 15

RESULT 22

YK01_PYRHO
 ID YK01_PYRHO STANDARD; PRT; 147 AA.
 AC 057781;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein PH2001.
 GN PH2001.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawanabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sakine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).

 This SWISS-PROT entry is copyright. It is produced through a coll.
 between the Swiss Institute of Bioinformatics and the EMBL out-
 station of the European Bioinformatics Institute. There are no restriction
 use by non-profit institutions as long as its content is in
 modified and this statement is not removed. Usage by and for com-
 mercial entities requires a license agreement (See <http://www.isb-sib.ch/>).

 CC EMBL; AP000001; BAA31940.1;
 CC EMBL; AP000007; BAA31940.1; JOINED.
 CC EMBL; AP000007; BAA31943.1;
 CC EMBL; AP000001; BAA31943.1; JOINED.
 CC PIR; A71217; A71217.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 41 61 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 SQ SEQUENCE 147 AA; 15324 MW; 247ED12FCE265B9 CRC64;

Query Match

2.5%; Score 7; DB 1; Length 147;

arity 100.0%; Pred. No. 38;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;

GLL 73
 ||||
 GLL 51

STANDARD; PRT; 150 AA.

Rel. 41, Created)
 Rel. 41, Last sequence update)
 Rel. 41, Last annotation update)
 eta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
 or ligand superfamily member 3) (Fragment).
 OR TNFC.
 '19).
 tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 heria; Cetartiodactyla; Suina; Suidae; Sus.
 '23;
 I N.A.
 white; TISSUE-Fibroblast;
 '615; PubMed=11169259;
 logel-Gallard C., Cattolico L., Duprat S., Vaiman M.,
 the swine major histocompatibility complex region
 l non-classical class I genes.";
 ns 57:55-65(2001).
 Cytokine that binds to LTBR/TNFRSF3. May play a specific
 immune response regulation. Provides the membrane anchor
 attachment of the heterotrimeric complex to the cell

Heterotrimer of either two LTB and one LTA subunits or
 valent) two LTA and one LTB subunits.
 AR LOCATION: Type II membrane protein (By similarity).
 Y: Belongs to the tumor necrosis factor family.

OT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL Outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 aires a license agreement (See <http://www.isb-sib.ch/announce/>
 ail to license@isb-sib.ch).

4; CAB63851.1; -;
 006053; TNF abc.
 006052; TNF family.
 008983; TNF like.
 003636; TNF_subf.
 ; TNF; 1.
 34; TNFCROISFCT.
 012; TNF_subf; 1.
 7; TNF; 1.
 251; TNF 1; 1.
 049; TNF 2; 1.
 coprotein.
 1
 28 128
 0 AA; 16423 MW; FESC4CC657658B48 CRC64;

2.5%; Score 7; DB 1; Length 150;
 arity 100.0%; Pred. No. 38;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;

LYC 200
 ||||
 LYC 43

RESULT 24

RA05_ORYSA STANDARD; PRT; 157 AA.

AC Q01881;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Seed allergenic protein RA5 precursor.
 GN RA5.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyt;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed;
 RX MEDLINE=93144699; PubMed=7678765;
 RA Adachi T., Izumi H., Yamada T., Tanaka K., Takeuchi S.,
 RA Nakamura R., Matsuoka T.;
 RT "Gene structure and expression of rice seed allergenic proteins
 belonging to the alpha-amylase/trypsin inhibitor family.";
 RL Plant Mol. Biol. 21:239-248(1993).
 CC -!- PTM: Five disulfide bonds are present (By similarity).
 CC -!- ALLERGEN: Causes an allergic reaction in human.
 CC -!- SIMILARITY: Belongs to the cereal trypsin/alpha-amylase inhib
 family.

CC This SWISS-PROT entry is copyright. It is produced through a coll
 between the Swiss Institute of Bioinformatics and the EMBL Out
 the European Bioinformatics Institute. There are no restriction
 use by non-profit institutions as long as its content is ir
 modified and this statement is not removed. Usage by and for c
 entities requires a license agreement (See <http://www.isb-sib.ch/>
 or send an email to license@isb-sib.ch).

CC EMBL; D11430; BAA01996.1; -;
 DR PIR; S31078; S31078.
 DR HSSP; P01085; 1HSS.
 DR Gramene; Q01881; -;
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR006106; Amylase_inhib.
 DR InterPro; IPR006105; Try/amyl_inhib.
 DR Pfam; PF00234; try_alpha_amyl; 1.
 DR PRINTS; PR00808; AMLASINBTR.
 DR SMART; SM00499; AAI; 1.
 DR PROSITE; PS00426; CEREAL_TRYP_AMYL_INH; 1.
 KW Allergen; Multigene family; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 157 SEED ALLERGENIC PROTEIN RA5;
 SQ SEQUENCE 157 AA; 17118 MW; C8A5495FPB399B6 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 157;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G

QY 72 LLLAVWS 78
 |||||
 DB 11 LLLAVWS 17

RESULT 25

HLPA_ECOLI STANDARD; PRT; 161 AA.

AC P11457;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Histone-like protein HLP-1 precursor (DNA-binding 17 kDa protein)
 GN HLPA OR SKP OR OMPH OR B0178 OR C0215 OR Z0190 OR ECS0180 OR SF01
 GN OR S0171.
 OS Escherichia coli, O6,
 OS Escherichia coli

D90651.
 DNEC17.
 455; hlpA.
 005632; Omph.
 ; Omph; 1.
 Outer membrane; Signal; Complete proteome.
 1 20
 21 161 HISTONE-LIKE PROTEIN HLP-1.
 1 AA; 17688 MW; 2A966BBD83F3E675 CRC64;
 2.5%; Score 7; DB 1; Length 161;
 arity 100.0%; Pred. No. 41;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ALA 68
 ||||
 ALA 15
 STANDARD; PRT; 170 AA.
 Rel. 28, Created)
 Rel. 28, Last sequence update)
 Rel. 41, Last annotation update)
 protein D2007.4 in chromosome III.
 s elegans.
 Razos; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 Peleiderinae; Caenorhabditis.
 39;
 I.N.A.
 I.N2.
 718; PubMed=7906398;
 nsough R., Anderson K., Baynes C., Berks M.,
 Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 Jones M., Kershaw J., Kirsten J., Laister N.,
 Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
 Don N., Smith A., Smith M., Sonhammer E., Staden R.,
 Hierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 Watson A., Weinstock L., Wilkinson-Sproat J.,
 ntiguous nucleotide sequence from chromosome III of C.
 38(1994).
 OT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL out-
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 fires a license agreement (See <http://www.isb-sib.ch/announce/>
 mail to license@isb-sib.ch).

 AAA27999.1; -.
 S44789.
 7.4; CE00129.
 protein.
 70 AA; 19396 MW; 22301D7C65638135 CRC64;
 2.5%; Score 7; DB 1; Length 170;
 arity 100.0%; Pred. No. 43;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RCL 227
 ||||
 RCL 100
 RESULT 27
 LACB_BUBBU STANDARD; PRT; 180 AA.
 ID LACB_BUBBU
 AC P02755; O62822;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Beta-lactoglobulin precursor (Beta-LG).
 GN LGB.
 OS Bubalus bubalis (Domestic water buffalo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bubalus.
 OX NCBI_TaxID=89462;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=99304500; PubMed=10376212;
 RA Das P., Jain S., Nayak S., Apparaio K.B.C., Totey S.M., Garg L.C.;
 RT "Molecular Cloning and sequence analysis of the cDNA encoding
 beta-lactoglobulin in Bubalus bubalis.";
 RL DNA Seq. 10:105-108 (1999).
 RN [2]
 RP SEQUENCE OF 19-180.
 RA Kolde H.-J., Liberatori J., Braunitzer G.;
 RT "The amino acid sequence of the water buffalo beta-lactoglobulin
 Milchwissenschaft 36:83-86 (1981).
 RL
 CC -!- FUNCTION: Primary component of whey, it binds retinol and is
 probably involved in the transport of that molecule.
 CC -!- SUBUNIT: Under physiological conditions beta-lactoglobulin e;
 as an equilibrium mixture of monomeric and dimeric forms.
 CC -!- SURCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Synthesized in mammary gland and secret
 in milk.
 CC -!- PTM: Alternate disulfide bonds occur in equal amounts.
 CC -!- SIMILARITY: Belongs to the lipocalin family.

 CC This SWISS-PROT entry is copyright. It is produced through a col-
 between the Swiss Institute of Bioinformatics and the EMBL out-
 the European Bioinformatics Institute. There are no restriction
 use by non-profit institutions as long as its content is i;
 modified and this statement is not removed. Usage by and for
 entities requires a license agreement (See <http://www.isb-sib.ch/>
 or send an email to license@isb-sib.ch).

 EMBL; AJ005429; CAA06532.1; -.
 InterPro; IPR002345; Lipocalin.
 DR InterPro; IPR000566; Lipocalin_cytfabp.
 DR Pfam; PF00061; lipocalin; 1.
 DR PRINTS; PR00179; LIPOCALIN.
 DR PROSITE; PS00213; LIPOCALIN; 1.
 KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.
 FT SIGNAL
 1 18
 FT CHAIN 19 180 BETA-LACTOGLOBULIN.
 FT DISULFID 84 178
 FT DISULFID 124 137
 FT DISULFID 124 139
 FT DISULFID 124 139
 SQ SEQUENCE 180 AA; 20223 MW; 6836C97B2C2E33CF CRC64;
 ALTERNATE.
 Query Match 2.5%; Score 7; DB 1; Length 180;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 QY 63 LGLALAC 69
 |||||
 DB 8 LGLALAC 14
 RESULT 28
 LACB_CAPHI STANDARD; PRT; 180 AA.
 ID LACB_CAPHI
 AC P02756;
 2.5%; Score 7; DB 1; Length 180;
 arity 100.0%; Pred. No. 45;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RCL 227
 |||||
 RCL 100

rel. 01, Created)
 rel. 22, Last sequence update)
 rel. 43, Last annotation update)
 nulin precursor (Beta-LG).
 (Goat).
 azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 eria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 nae; Capra.
 5;
 N.A.
 agrus; TISSUE=Mammary gland;
 59; PubMed=8226387;
 ill A., Sanchez A.;
 sequencing of the cDNA encoding goat beta-
 .";
 71:2832-2832(1993).
 N.A.
 Kim J., Yu M.;
 I-1993) to the EMBL/GenBank/DBJ databases.
 N.A.
 51; PubMed=7699130;
 A., Sanchez A.;
 nence of the caprine beta-lactoglobulin gene.";
 77:3493-3497(1994).
 1-180.
 11; PubMed=511095;
 unitzer G., Schrank B., Stangl A.;
 d sequence of goat beta-lactoglobulin.";
 2. Physiol. Chem. 360:1595-1604(1979).
 Primary component of whey, it binds retinol and is
 involved in the transport of that molecule.
 Under physiological conditions beta-lactoglobulin exists
 librium mixture of monomeric and dimeric forms.
 AR LOCATION: Secreted.
 ICIFICITY: Synthesized in mammary gland and secreted
 nate disulfide bonds occur in equal amounts.
 : Belongs to the lipocalin family.
 NT entry is copyright. It is produced through a collaboration
 wiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 his statement is not removed. Usage by and for commercial
 res a license agreement (see <http://www.isb-sib.ch/announce/>
 il to license@isb-sib.ch).

 CAA41385.1; -;
 CAA79623.1; -;
 CAA79624.1; -;
 CAA83946.1; -;
 AGT.
 1B00.
 002345; Lipocalin.
 00566; Lipocalin_cyFABP.
 : lipocalin; 1.
 79; LIPOCALIN.
 213; LIPOCALIN; 1.
 etinol-binding; Transport; Lipocalin; Signal.
 1 18
 19 180 BETA-LACTOGLOBULIN.
 34 178
 24 137
 24 139
) AA; 19975 MW; C2449BH02A1A80F1 CRC64;
 2.5%; Score 7; DB 1; Length 180;

Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 QY 63 LGLALAC 69
 Db 8 LGLALAC 14
 RESULT 29
 LACB SHEEP
 ID LACB SHEEP STANDARD; PRT; 180 AA.
 AC P02757;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Beta-lactoglobulin 1/B, 2/A, and 3/C precursor.
 OS Ovis aries (Sheep), and
 OS Ovis orientalis musimon (Mouflon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940, 9938;
 RN [1]
 RP SEQUENCE FROM N.A. (BLG 1 AND 2).
 RC SPECIES=Sheep;
 RX MEDLINE=88172489; PubMed=3351935;
 RA Ali S., Clark A.J.;
 RT "Characterization of the gene encoding ovine beta-lactoglobulin.
 RT Similarity to the genes for retinol binding protein and other
 RT secretory proteins.";
 RL J. Mol. Biol. 199:415-426(1988).
 RN [2]
 RP SEQUENCE FROM N.A. (BLG 1).
 RC SPECIES=Sheep;
 RX MEDLINE=87049827; PubMed=3096387;
 RA Gaye P., Hue-Delahaie D., Mercier J.-C., Soulier S., Vilotte J.-L
 RA Furet J.-P.;
 RT "Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and
 RT levels during functional differentiation of the mammary gland.";
 RL Biochimie 68:1097-1107(1986).
 RN [3]
 RP SEQUENCE FROM N.A. (BLG 1).
 RC SPECIES=Sheep;
 RX MEDLINE=89057492; PubMed=3194215;
 RA Harris S., Ali S., Anderson S., Archibald A.L., Clark A.J.;
 RT "Complete nucleotide sequence of the genomic ovine beta-lactoglob
 RT gene.";
 RL Nucleic Acids Res. 16:10379-10380(1988).
 RN [4]
 RP SEQUENCE FROM N.A. (BLG 1 AND 2).
 RC SPECIES=Sheep;
 RX MEDLINE=91007276; PubMed=1976573;
 RA Ali S., McClenaghan M., Simons J.P., Clark A.J.;
 RT "Characterisation of the alleles encoding ovine beta-lactoglobuli;
 RT Gene 91:201-207(1990).
 RN [5]
 RP SEQUENCE OF 19-180 (BLG 2).
 RC SPECIES=Sheep;
 RX MEDLINE=80219294; PubMed=6155855;
 RA Preaux G., Braunitzer G., Kolde H.-J.;
 RT "Primary structure of ovine beta-lactoglobulin.";
 RL Arch. Int. Physiol. Biochim. 88:845-846(1980).
 RN [6]
 RP SEQUENCE OF 19-180 (BLG 3).
 RC SPECIES=Sheep;
 RX MEDLINE=89374823; PubMed=2775495;
 RA Erhardt G., Godovac-Zimmermann J., Conti A.;
 RT "Isolation and complete primary sequence of a new ovine wild-type
 RT beta-lactoglobulin C.";
 RL Biol. Chem. Hoppe-Seyler 370:757-762(1989).
 RN [7]
 RP SEQUENCE OF 19-180 (BLG B).

```

musimon;
1996; PubMed=3426802;
1996; J. Conti A., Napolitano L.;
1996; amino-acid sequence of dimeric beta-lactoglobulin from
1996; amnon musimon) milk.;
1996; Saylor 368:1313-1319(1987).
1996; LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
1996; INOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
1996; ECULE.
1996; Under physiological conditions beta-lactoglobulin exists
1996; in a mixture of monomeric and dimeric forms.
1996; disulfide bonds occur in equal amounts.
1996; belongs to the lipocalin family.
1996; ROT entry is copyright. It is produced through a collaboration
1996; Swiss Institute of Bioinformatics and the EMBL outstation -
1996; Bioinformatics Institute. There are no restrictions on its
1996; profit institutions as long as its content is in no way
1996; this statement is not removed. Usage by and for commercial
1996; a license agreement (see http://www.isb-sib.ch/announce/
1996; mail to license@isb-sib.ch).
1996; CAA28204.1; -
1996; CAA31305.1; -
1996; CAA30059.1; ALT SEQ.
1996; CAA30059.1; JOINED.
1996; CAA30059.1; JOINED.
1996; CAA30059.1; JOINED.
1996; CAA30059.1; JOINED.
1996; CAA30059.1; JOINED.
1996; AAA31510.1; -
1996; AAA31510.1; JOINED.
1996; AAA31510.1; JOINED.
1996; AAA31510.1; JOINED.
1996; LGSB.
1996; IBSQ.
1996; R002345; Lipocalin.
1996; R000566; Lipocalin_cytFABP.
1996; 1; lipocalin; 1.
1996; 179; LIPOCALIN.
1996; 0213; LIPOCALIN; 1.
1996; Retinol-binding; Transport; Signal; Lipocalin.
1996; 1 18
1996; 19 180 BETA-LACTOGLOBULIN.
1996; 84 178
1996; 124 137
1996; 124 139
1996; 38 138
1996; 166 166
1996; 30 AA; 19921 MW; BABCB289E757333 CRC64;
1996; 2.5%; Score 7; DB 1; Length 180;
1996; larity 100.0%; Pred. No. 45;
1996; conservative 0; Mismatches 0; Indels 0; Gaps 0;
1996; ALAC 69
1996; |||||
1996; ALAC 14
1996; STANDARD; PRT; 181 AA.
1996; (Rel. 42, Created)
1996; (Rel. 42, Last sequence update)
1996; (Rel. 42, Last annotation update)
1996; phorbosyltransferase (EC 2.4.2.7) (APRT).
1996; 2.
1996; neidensis.
1996; oteobacteria; Gammaproteobacteria; Alteromonadales;
1996; ceae; Shewanella.
1996; 0863;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.F., Nelson K.E., Gaidos E.J., Nelson R.
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.
RA Meyer T., Tsapin A., Scott J., Bearan M., Brinkac L., Daugherty J.
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Uitterback T.R., McDonald L.A., C.
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.
RT "Genome sequence of the dissimilatory metal ion-reducing bacteri
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- FUNCTION: Catalyzes a salvage reaction resulting in the form
CC of AMP, that is energetically less costly than de novo synthe
CC -!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC D-ribose 1-diphosphate.
CC -!- PATHWAY: Purine salvage.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the purine/pyrimidine
CC phosphoribosyltransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a col
CC between the Swiss Institute of Bioinformatics and the EMBL ou
CC the European Bioinformatics Institute. There are no restrictio
CC use by non-profit institutions as long as its content is i
CC modified and this statement is not removed. Usage by and for
CC entities requires a license agreement (See http://www.isb-sib.ch
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE015643; AAN55062.1; -
CC TIGR; SO2012; -
CC HAMAP; MF_00004; -; 1.
CC InterPro; IPR005764; Ade phspho trans.
CC InterPro; IPR002375; Pr/py rp trans.
CC InterPro; IPR000836; Prtransferase.
CC Pfam; PF00156; Pribosyltran; 1.
CC TIGRFAMs; TIGR01090; apt; 1.
CC PROSITE; PS00103; PUR PYR PR TRANSFER; 1.
CC TRANSFERASE; Glycosyltransferase; Purine salvage; Complete prote
CC SEQUENCE 181 AA; 19543 MW; C4255A59C4632CA4 CRC64;
CC -----
CC Query Match 2.5%; Score 7; DB 1; Length 181;
CC Best Local Similarity 100.0%; Pred. No. 45;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0;
CC QY 58 PLALGIG 64
CC |||||
CC Db 71 PLALGIG 77
CC -----
CC RESULT 31
CC DEF1 BIFLO
CC ID DEF1 BIFLO STANDARD; PRT; 217 AA.
CC AC Q8G534;
CC DT 15-MAR-2004 (Rel. 43, Created)
CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Peptide deformylase 1 (EC 3.5.1.88) (PDF 1) (Polypeptide deformy
CC 1).
CC GN DEF1 OR BL1186.
CC OS Bifidobacterium longum.
CC OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
CC OC Bifidobacteriaceae; Bifidobacterium.
CC NCBI_TaxID=216816;
CC [1]
CC RN SEQUENCE FROM N.A.
CC STRAIN=NCC 2705;
CC MEDLINE=22294977; PubMed=12381787;
CC Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,

```

Allen M.-C., Desiere F., Bork P., Delley M.,
 Arigoni F.,
 sequence of Bifidobacterium longum reflects its adaptation
 to the gastrointestinal tract.";
 J. Biol. Chem. 275:14422-14427 (2000).
 2. Removes the formyl group from the N-terminal Met of
 ribosomal proteins. Requires at least a dipeptide for an
 active site. N-terminal L-methionine is a
 rate of reaction. The enzyme has broad specificity at
 the formyl group but the enzyme has broad specificity at
 the N-terminal L-methionine.
 (By similarity).
 ACTIVITY: Formyl-L-methionyl peptide + H₂O = formate +
 peptide.
 Binds 1 iron(II) ion (By similarity).
 Y: Belongs to the polypeptide deformylase family.
 OT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL
 Outstation. There are no restrictions on its
 use by non-profit institutions as long as its content is
 not removed. Usage by and for commercial
 purposes requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 3; AAN24991.1; -;
 63; -; 1.
 000181; Fmet deformylase.
 1; Pep deformylase.
 76; PDEFORMYLASE.
 844; Pep deformylase; 1.
 nthesis; Hydrolyase; Iron; Complete proteome.
 72 172 BY SIMILARITY.
 29 129 IRON (BY SIMILARITY).
 71 171 IRON (BY SIMILARITY).
 75 175 IRON (BY SIMILARITY).
 7 AA; 24443 MW; 52F5B469B6F47163 CRC64;
 2.5%; Score 7; DB 1; Length 217;
 identity 100.0%; Pred. No. 53;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ALA 68
 |||||
 ALA 86
 STANDARD; PRT; 230 AA.
 Rel. 41, Created
 Rel. 41, Last sequence update
 Rel. 42, Last annotation update
 Sport complex protein rnfE.
 54 OR STY1668 OR T1322.
 phimurium, and
 phi.
 teobacteria; Gammaproteobacteria; Enterobacteriales;
 aceae; Salmonella.
 2, 601;
 N.A.
 himurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 948; PubMed=11677609;
 Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 Guyon C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 H., Florea L., Miller W., Stoneking T., Nhan M.,
 Wilson R.K.,
 some sequence of Salmonella enterica serovar Typhimurium
 2-856(2001).
 N.A.

SPECIES=S.typhi; STRAIN=CT18;
 MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wai
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.
 RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RT Nature 413:848-852(2001).
 RL [3]
 RP SEQUENCE FROM N.A.
 RP SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RX Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strain
 RT and CT18";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -!- FUNCTION: May be part of a membrane complex involved in elect
 CC transport (By similarity).
 CC -!- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,
 CC rnfD, rnfE and rnfG (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membra
 CC (Potential).
 CC -!- SIMILARITY: Belongs to the nqrDE/rnfAE family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a coll
 CC between the Swiss Institute of Bioinformatics and the EMBL out
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is ir
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE008763; AAL20376.1; -;
 CC EMBL; AL627271; CAD01913.1; -;
 CC EMBL; AE016838; AAC68972.1; -;
 CC StyGene; SG????; rnfE.
 CC HAMAP; MF_00478; -; 1.
 CC InterPro; IPR003667; Rnf_Nqr.
 CC Pfam; PF02508; Rnf_Nqr; I.
 KW Electron transport; Transmembrane; Inner membrane; Complete prote
 FT TRANSMEM 34 56 POTENTIAL.
 FT TRANSMEM 69 87 POTENTIAL.
 FT TRANSMEM 91 113 POTENTIAL.
 FT TRANSMEM 126 148 POTENTIAL.
 FT TRANSMEM 183 205 POTENTIAL.
 SQ SEQUENCE 230 AA; 24318 MW; E198B4CEA13F249E CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 230;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; C
 QY 60 ALGGLA 66
 |||||
 DB 38 ALGGLA 44
 RESULT 33
 RNFE_ECO57
 ID RNFE_ECO57 STANDARD; PRT; 231 AA.
 AC P58344;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Electron transport complex protein rnfE.
 GN RNFE OR Z2642 OR ECS2341.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

aceae; Escherichia.
 1334;
 4 N.A.
 17 / BDL933 / ATCC 700927;
 1935; PubMed11206551;
 Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 yehew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 ickett J., Klink S., Boutin A., Shao Y., Miller L.,
 . Davis M.W., Lim A., Dimalanta E.T., Potamoulis K.,
 Mantharman T.S., Lin J., Yen G., Schwartz D.C.,
 Blattner F.R.;
 nce of enterohaemorrhagic Escherichia coli O157:H7.;
 29-533 (2001).
 4 N.A.
 17 / RMD 050952;
 5231; PubMed11258796;
 akino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 itsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 mi H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 uba T., Hattori M., Shinagawa H.;
 ome sequence of enterohemorrhagic Escherichia coli
 genomic comparison with a laboratory strain K-12.;
 [-22(2001).
 : May be part of a membrane complex involved in electron
 : (By similarity).
 : Composed of at least six subunits; rnfA, rnfB, rnfC,
 E and rnfG (By similarity).
 AR LOCATION: Integral membrane protein. Inner membrane
 IV).
 IV): Belongs to the nqrDE/rnfAE family.
 ROT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 -profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 izes a license agreement (See <http://www.isb-sib.ch/announce/>
 mail to license@isb-sib.ch).
 36; AAG56621.1; -;
 58; BAB35764.1; -;
 A85770.
 E90921.
 178: -; 1.
 3003667; Rnf Nqr.
 3; Rnf-Nqr; 1.
 isport; Transmembrane; Inner membrane; Complete proteome.
 1 38 PERIPLASMIC (POTENTIAL).
 39 59 POTENTIAL.
 60 62 CYTOPLASMIC (POTENTIAL).
 63 83 POTENTIAL.
 84 85 PERIPLASMIC (POTENTIAL).
 86 106 POTENTIAL.
 107 124 CYTOPLASMIC (POTENTIAL).
 125 145 POTENTIAL.
 146 181 PERIPLASMIC (POTENTIAL).
 182 202 POTENTIAL.
 203 231 CYTOPLASMIC (POTENTIAL).
 31 AA; 24489 MW; D4A2CA2D292604C3 CRC64;
 2.5%; Score 7; DB 1; Length 231;
 larity 100.0%; Pred. No. 56;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 LGLA 66
 ||||
 LGLA 44
 ID RNFE ECOLI STANDARD; PRT; 231 AA.
 AC P77179;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Electron transport complex protein rnfE.
 GN RNFE OR B1632.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
 RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
 RA Oshima T., Saito N., Sanpei G., Seki Y., Sivasubraman S.,
 RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
 RA Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 RN [3]
 RP TOPOLOGY.
 RC STRAIN=K12 / JM109;
 RX MEDLINE=99342054; PubMed=10411911;
 RA Saeef A., Johansson M., Wallin E., von Heijne G.;
 RT "Divergent evolution of membrane protein topology: the Escherich
 coli RnfA and RnfE homologues";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:8540-8544(1999).
 CC -!- FUNCTION: May be part of a membrane complex involved in elec
 transport (By similarity).
 CC -!- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC
 rnfD, rnfE and rnfG (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membr
 (Potential).
 CC -!- SIMILARITY: Belongs to the nqrDE/rnfAE family.
 CC This SWISS-PROT entry is copyright. It is produced through a col
 between the Swiss Institute of Bioinformatics and the EMBL ou
 the European Bioinformatics Institute. There are no restrictio
 use by non-profit institutions as long as its content is i
 modified and this statement is not removed. Usage by and for
 entities requires a license agreement (See <http://www.isb-sib.ch>
 or send an email to license@isb-sib.ch).
 CC EMBL; AE000258; AAC74704.1; -;
 DR EMBL; D90806; BAA15386.1; -;
 DR EMBL; D90807; BAA15393.1; -;
 DR EMBL; D90808; BAA15416.1; -;
 DR PIR; B64920; B64920.
 DR EcoGene; EGI13938; rnfE.
 DR HAMAP; MF_00478; -; 1.
 DR InterPro; IPR003667; Rnf_Nqr.
 DR Pfam; PF02508; Rnf-Nqr; 1.
 DR Electron transport; Transmembrane; Inner membrane; Complete prot.
 KW DOMAIN 1 38 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 39 59 POTENTIAL.
 FT DOMAIN 60 62 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 63 83 POTENTIAL.

4 85 PERIPLASMIC (POTENTIAL).
6 106 POTENTIAL.
7 124 CYTOPLASMIC (POTENTIAL).
5 145 POTENTIAL.
6 181 PERIPLASMIC (POTENTIAL).
2 202 POTENTIAL.
3 231 CYTOPLASMIC (POTENTIAL).
AA; 24459 MW; CFA37A2D292604C3 CRC64;
2.5%; Score 7; DB 1; Length 231;
urity 100.0%; Pred.No.56;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
ILA 66
||
ILA 44
STANDARD; PRT; 233 AA.
rel. 34, Created)
rel. 42, Last sequence update)
rel. 42, Last annotation update)
III (EC 3.1.26.3) (RNase III).
stii.
eobacteria; Gammaproteobacteria; Legionellales;
Coxiella.
;
N.A.
/;
751; PubMed=7830573;
er T.A., Powell B.S., Court D.L.;
he rnc locus of Coxiella burnetii.";
l. 14:291-300(1994).
N.A.
ile phase I / RSA 493;
557; PubMed=12704232;
paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
ard N.L., Tettelin H., Daviden T.M., Beanan M.J.,
ugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
ee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
Samuel J.E., Fraser C.M., Heidelberg J.F.;
me sequence of the Q-fever pathogen, Coxiella
ad. Sci. U.S.A. 100:5455-5460(2003).
Digests double-stranded RNA. Involved in the processing
nal RNA precursors and of some mRNAs (By similarity).
ACTIVITY: Endonucleolytic cleavage to 5'-
noster.
AR LOCATION: Cytoplasmic.
Y: Contains 1 DBRM (double-stranded RNA-binding) domain.
Y: Contains 1 RNase III domain.
JT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ires a license agreement (See <http://www.isb-sib.ch/announce/>
ail to license@isb-sib.ch).
AAA69690.1; -.
4; AA091000.1; -.
S60767.
04; -; 1.
001159; DS RBD.
000999; RNase_III.

DR Pfam; PF00035; dsrm; 1.
DR PF00636; Ribonuclease_3; 1.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 1.
DR PROSITE; PS0137; DS RBD; 1.
DR PROSITE; PS00517; RNase_3_1; 1.
DR PROSITE; PS0142; RNase_3_2; 1.
KW Hydrolase; Nuclease; Endonuclease; RNA-binding; Complete proteome
FT DOMAIN 4 126 RNase III.
FT DOMAIN 204 220 DBRM.
FT CONFLICT 116 116 A -> T (IN REF. 1).
SQ SEQUENCE 233 AA; 26199 MW; 1A11CB5FD960784F CRC64;
Query Match 2.5%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred.No.56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G
QY 10 ARRLPLP 16
|||
Db 164 ARRLPLP 170
RESULT 36
RNFE YERPE STANDARD; PRT; 233 AA.
AC Q8ZED4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNFE OR YPO2240 OR Y2081.
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Tithall R.W., Holden M.T.G.
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Versinia pestis, the causative agent of plagu
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Y
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz J
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Pla
RA Straley S.C., McDonough K.A., Nilles M.L., Matso'
RA Perry R.D.;
RT "Genome sequence of Versinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
CC -!- FUNCTION: May be part of a membrane comple
transport (By similarity).
CC -!- SUBUNIT: Composed of at least six subunit
rnfD, rnfE and rnfG (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane
potential).
CC -!- SIMILARITY: Belongs to the nqrDE/rnfAE /
This SWISS-PROT entry is copyright. It is y
between the Swiss Institute of Bioinforma
the European Bioinformatics Institute. Th
use by non-profit institutions as long
modified and this statement is not remov
entities requires a license agreement (S

mail to license@isb-sib.ch).

51; CAC91046.1; -;
10; AAM85645.1; -;
AB0273.
478; -; 1.
R003667; Rnf Nqr.
3; Rnf-Nqr; 1.
asport; Transmembrane; Inner membrane; Complete proteome.
34 56 POTENTIAL.
69 87 POTENTIAL.
91 113 POTENTIAL.
126 148 POTENTIAL.
184 206 POTENTIAL.
33 AA; 24587 MW; 491E18F335B8CB90 CRC64;

2.5%; Score 7; DB 1; Length 233;
Larity 100.0%; Pred. No. 56;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LGLA 66
|||||
LGLA 44

46; STANDARD; PRT; 235 AA.

(Rel. 35, Created)
(Rel. 35, Last sequence update)
(Rel. 41, Last annotation update)
nsport complex protein rnfE.

88.
influenzae.
oteobacteria; Gammaproteobacteria; Pasteurellales;
eae; Haemophilus.
27;

M N.A.
KW20 / ATCC 51907;
0630; PubMed=7542800;
R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
Phillips C.A., Spriggs T., Hedblom E., Corton M.D.,
R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
Ritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

e random sequencing and assembly of Haemophilus influenzae
496-512(1995).

: May be part of a membrane complex involved in electron
t (By similarity).

Composed of at least six subunits; rnfA, rnfB, rnfC,
rnfE and rnfG (By similarity).

LAR LOCATION: Integral membrane protein. Inner membrane
al).

TY: Belongs to the nqrDE/rnfAE family.

ROT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
-profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
quires a license agreement (See <http://www.isb-sib.ch/announce/>
mail to license@isb-sib.ch).

; AAC23334.1; -;
; I64174.
; -;

DR HAMAP; MF_00478; -; 1.
DR InterPro; IPR003667; Rnf Nqr.
DR Pfam; PF02508; Rnf-Nqr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete prot
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
SQ SEQUENCE 235 AA; 25845 MW; C054FE596647837A CRC64;

Query Match 2.5%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 60 ALGLGLA 66
Db 62 ALGLGLA 68

RESULT 38

TN14 MOUSE STANDARD; PRT; 239 AA.

ID TN14 MOUSE
AC QSOYH9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 14.
GN TNFSF14 OR LIGHT.
OS Mus musculus (Mouse).
OC Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mu
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20165223; PubMed=10700230;
RA Tamada K., Shimozaaki K., Chapoval A.I., Zhu G., Sica G., Flies I
RA Boone T., Hsu H., Fu Y.-X., Nagata S., Ni J., Chen L.;
RT "Modulation of T-cell-mediated immunity in tumor and graft-versu
RT disease models through the LIGHT co-stimulatory pathway.";
RL Nat. Med. 6:283-289(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Fetal liver;
RC MEDLINE=20354998; PubMed=10894944;
RA Misawa K., Nosaka T., Kojima T., Hirai M., Kitamura T.;
RT "Molecular cloning and characterization of a mouse homolog of hu
RT TNFSF14, a member of the TNF superfamily.";
RL Cytogenet. Cell Genet. 89:89-91(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymphoma;
RC Force W.R., Todd P.K., Mikayama T.;
RT "Mouse LIGHT; molecular genetics, ligand binding and expression.
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to th
decoy receptor TNFRSF6B modulates its effects. Activates NFK
and stimulates the proliferation of T cells.
-1- SUBUNIT: Homotrimer (By similarity).
-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted
similarity).
-1- PTM: The soluble form derives from the membrane form by
proteolytic processing.
-1- SIMILARITY: Belongs to the tumor necrosis factor family.
CC This SWISS-PROT entry is copyright. It is produced through a col
between the Swiss Institute of Bioinformatics and the EMBL ou
the European Bioinformatics Institute. There are no restrictio
use by non-profit institutions as long as its content is i
modified and this statement is not removed. Usage by and for
entities requires a license agreement (See <http://www.isb-sib.ch>
or send an email to license@isb-sib.ch).

5; AAF76453.1; -;
 5; BAA8559.1; -;
 3; AAF36722.1; -;
 4TSV.
 317; Tnfstf14.
 006053; TNF_abi.
 006052; TNF_family.
 009883; TNF_like.
 003636; TNF_subf.
 ; TNF; 1.
 34; TNECROSISFCT.
 012; TNF_subf; 1.
 7; TNF; 1.
 251; TNF_1; FALSE_NEG.
 049; TNF_2; 1.
 asmembrane; Glycoprotein; Signal-anchor.
 1 239
 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
 MEMBER 14, MEMBRANE FORM.
 32 239
 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
 MEMBER 14, SOLUBLE FORM.
 1 37
 CYTOPLASMIC (POTENTIAL).
 38 58
 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 (POTENTIAL).
 59 239
 EXTRACELLULAR (POTENTIAL).
 31 82
 CLEAVAGE (POTENTIAL).
 52 187
 POTENTIAL.
 30 100
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 31 191
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 3 AA; 26338 MW; 217874AC71AD6BE3 CRC64;
 2.5%; Score 7; DB 1; Length 239;
 arity 100.0%; Pred.No. 58;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RPR 127
 ||||
 RPR 227
 STANDARD; PRT; 240 AA.
 rel. 41, Created)
 rel. 41, Last sequence update)
 rel. 41, Last annotation update)
 sport complex protein rnfE.
 4.
 aruginosa.
 ceobacteria; Gammaproteobacteria; Pseudomonadales;
 eae; Pseudomonas.
 7;
 N.A.
 5692 / PA01;
 337; PubMed=10984043;
 Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
 Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Soltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 Culter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 pence D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 ler M.H., Hancock R.E.W., Lory S., Olson M.V.;
 ome sequence of *Pseudomonas aeruginosa* PA01, an
 pathogen.";
 9-964(2000).
 May be part of a membrane complex involved in electron
 (By similarity).
 Composed of at least six subunits; rnfA, rnfB, rnfC,
 E and rnfG (By similarity).
 AR LOCATION: Integral membrane protein. Inner membrane
 1).
 Y: Belongs to the nqrDE/rnfAE family.

CC This SWISS-PROT entry is copyright. It is produced through a coll
 CC between the Swiss Institute of Bioinformatics and the EMBL out
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is in
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AEO04770; AAG06882.1; -;
 DR PIR; G83208; G83208.
 DR HAMAP; MF 00478; -; 1.
 DR InterPro; IPR003667; Rnf_Nqr.
 DR Pfam; PF02508; Rnf-Nqr; 1.
 KW Electron transport; Transmembrane; Inner membrane; Complete prote
 FT TRANSMEM 41 61 POTENTIAL.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 95 115 POTENTIAL.
 FT TRANSMEM 130 150 POTENTIAL.
 FT TRANSMEM 184 204 POTENTIAL.
 SQ SEQUENCE 240 AA; 3D90687ED462D8B2 CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred.No. 58;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 Qy 60 ALGLGLA 66
 |||||
 Db 40 ALGLGLA 46
 RESULT 40
 MOEB HAEIN
 ID MOEB HAEIN STANDARD; PRT; 243 AA.
 AC P45211;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Molybdopterin biosynthesis protein moeb.
 GN MOEB OR CHLN OR HII449.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C
 RA Fine L.D., Pritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influ*
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
 CC (MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE I
 CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTI
 CC FACTOR (MOAD) (BY SIMILARITY).
 CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a coll
 CC between the Swiss Institute of Bioinformatics and the EMBL out
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is in
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC -----


```

AAC23099.1; -.
C64124.
-
R009036; MoeB.
R007901; MoeZ MoeB.
R000205; NAD.BS.
R000594; Thif_domain.
?; MoeZ MoeB; 1.
?; Thif; 1.
factor biosynthesis; Complete proteome.
13 AA; 26996 MW; 218A3382A975BDBD CRC64;
2.5%; Score 7; DB 1; Length 243;
larity 100.0%; Pred.No.59;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
NR01185
|||||
NR0116
STANDARD; PRT; 244 AA.
70; Q99761;
(Rel. 29, Created)
(Rel. 43, Last annotation update)
(Rel. 43, Last annotation update)
beta (IT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
cor ligand superfamily member 3).
3 OR INFC.
(Human).
stazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
theria; Primates; Catarrhini; Hominidae; Homo.
506;
M N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
1;
3881; PubMed:7916655;
Ngan-Ek A., Lawton P., Demarinis J., Tizard R.,
ession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
beta, a novel member of the TNF family that forms a
complex with lymphotoxin on the cell surface.";
356(1993).
M N.A. (ISOFORMS 1 AND 2).
5965; PubMed:9299492;
Renard N., Charlot C., Bienvenu J., Coiffier B.,
ion of two lymphotoxin beta isoforms expressed in human
1 lines and non-Hodgkin's lymphomas.";
phys. Res. Commun. 238:273-276(1997).
M N.A. (ISOFORM 1).
Milner C.M., Campbell R.D.;
r of the immunoglobulin superfamily and a V-Apase G
amongst the predicted products of novel genes close to the
the human MHC.";
EP-1997) to the EMBL/GenBank/DBJ databases.
M N.A. (ISOFORM 1).
dan A., Qin S., Shaffer T., James R., Ratcliffe A.,
ickhoff R., Loretz C., Madan A., Dors M., Young J.,
od L.;
the human major histocompatibility complex class III
CT-1999) to the EMBL/GenBank/DBJ databases.
M N.A. (ISOFORM 1).
amiya G., Oka A., Inoko H.;
s 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
EP-1999) to the EMBL/GenBank/DBJ databases.

```

```

[6]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLU-70 AND
PRO-111.
RA Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS ARG-84 AND PHE-87.
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytokine that binds to LTBR/TNFRSF3. May play a sp
CC role in immune response regulation. Provides the membrane an
CC for the attachment of the heterotrimeric complex to the cell
CC surface. Isoform 2 is probably non-functional.
CC -!- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
CC (less prevalent) one LTB and two LTA subunits.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q06643-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q06643-2; Sequence=VSP_006441, VSP_006442;
CC -!- TISSUE SPECIFICITY: Spleen and thymus.
CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a col
CC between the Swiss Institute of Bioinformatics and the EMBL ou
CC the European Bioinformatics Institute. There are no restrictio
CC use by non-profit institutions as long as its content is i
CC modified and this statement is not removed. Usage by and for
CC entities requires a license agreement (See http://www.isb-sib.ch
CC or send an email to license@isb-sib.ch).
-----
EMBL; L11016; AAA99888.1; -
EMBL; U99922; AAC51769.1; -
EMBL; U79029; AAB37442.1; -
EMBL; L11015; AAA36191.1; -
EMBL; Y14768; CAA75069.1; -
EMBL; AF129756; RAD18089.1; -
EMBL; AF000505; BAB63395.1; -
EMBL; AY070219; AAL49954.1; -
EMBL; AY070219; AAL49955.1; -
EMBL; AY216497; AAO21134.1; -
PIR; A46066; A46066.
PIR; JCS645; JCS645.
HSSP; P01374; 1TNF.
Genew; HGNC:6711; LTB.
MIM; 600978; -.
GO; GO:0005102; F:receptor binding; TAS.
GO; GO:0015070; F:toxin activity; NAS.
GO; GO:0007267; P:cell-cell signaling; TAS.
GO; GO:0007165; P:signal transduction; TAS.
InterPro; IPR006053; TNF_abc
InterPro; IPR006052; TNF_family.
InterPro; IPR008983; TNF_like.
InterPro; IPR003636; TNF_subf.
Pfam; PF00229; TNF; 1.
PRINTS; PR01234; TNECROSISFCT.
ProDom; PD002012; TNF_subf; 1.
SMART; SM00207; TNF; 1.
PROSITE; PS00251; TNF_1; 1.
PROSITE; PS00049; TNF_2; 1.
KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor;
KW Alternative splicing; Polymorphism.
FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO
FT (POTENTIAL).
FT DOMAIN 49 244 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 222 242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPPLIC 53 77 GLVETADFGAQAQQGLGFKLPEE -> GLGFRS

```


my-protein.
 AR LOCATION: Cytoplasmic (By similarity).
 Y: Belongs to the L/F-transferase family.

 NOT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 hires a license agreement (See <http://www.isb-sib.ch/announce/>
 mail to license@isb-sib.ch).

 J1; AA041258.1; -;
 588; -; 1.
 3; Leu Phe trans; 1.
 3R00667; aat; 1.
 Acyltransferase; Complete proteome.
 19 AA; 27404 MW; E4584D8A164EDB54 CRC64;

 Identity 2.5%; Score 7; DB 1; Length 249;
 Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

 3QDG 156
 |||||
 3QDG 116

 STANDARD; PRT; 249 AA.
 (Rel. 12, Created)
 (Rel. 12, Last sequence update)
 (Rel. 42, Last annotation update)
 Biosynthesis protein moeb.
 OR B0826.
 coli.
 teobacteria; Gammaproteobacteria; Enterobacteriales;
 laceae; Escherichia.
 52;
 4 N.A.
 1906; PubMed-3045084;
 sai Y., Saito T.;
 sequencing of the Escherichia coli chlen operon involved
 arin biosynthesis.";
 . 170:4097-4102(1988).
 4 N.A.
 MG1655;
 5617; PubMed-9278503;
 Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 llado-Vides J., Glaser J.D., Rode C.K., Mayhew G.F.,
 avis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 Y.;
 genome sequence of Escherichia coli K-12.";
 1453-1474(1997).
 4 N.A.
 1202; PubMed-8905232;
 iba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 itagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 omura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
 eki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 iuchi T.;
 A sequence of the Escherichia coli K-12 genome
 37 to the 12.7-28.0 min region on the linkage map.";
 37-155(1996).
 INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR

(MOLYBDOPTERIN), NECESSARY FOR MOLYBDOENZYMES. PLAYS A ROLE
 ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERT
 FACTOR (MOAD).
 -!- PATHWAY: Molybdenum cofactor biosynthesis.
 -!- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a col-
 between the Swiss Institute of Bioinformatics and the EMBL ou-
 the European Bioinformatics Institute. There are no restrictio-
 use by non-profit institutions as long as its content is i-
 modified and this statement is not removed. Usage by and for
 entities requires a license agreement (See <http://www.isb-sib.ch>
 or send an email to license@isb-sib.ch).

 EMBL; M21151; AAA23580.1; -;
 EMBL; AE000185; AAC73913.1; -;
 EMBL; D90720; BAA35514.1; -;
 EMBL; D90721; BAA35521.1; -;
 PIR; B32352; B32352.
 PDB; 1JW9; 21-NOV-01.
 PDB; 1JWA; 21-NOV-01.
 PDB; 1JWB; 21-NOV-01.
 EcoGene; EGI0154; moeb.
 InterPro; IPR009036; Moeb_Moeb.
 InterPro; IPR007901; Moeb_Moeb.
 InterPro; IPR000205; NAD_BS.
 InterPro; IPR000594; Thif_domain.
 Pfam; PF05237; Moeb_Moeb; 1.
 Pfam; PF00899; Thif; 1.
 Molybdenum cofactor biosynthesis; Complete proteome; 3D-structur
 KW Molybdenum cofactor biosynthesis; Complete proteome; 3D-structur
 SQ SEQUENCE 249 AA; 26719 MW; 12C77082B3F39D7D CRC64;

 Query Match 2.5%; Score 7; DB 1; Length 249;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 7; Conservativity 0; Mismatches 0; Indels 0;

 QY 179 LRYNRQI 185
 |||||
 Db 10 LRYNRQI 16

 RESULT 45
 MOEB_SALTY
 ID MOEB_SALTY STANDARD; PRT; 249 AA.
 AC Q56067;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Molybdopterin biosynthesis protein moeb.
 GN MOEB OR STM0845.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]_TaxID=602;
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RC Wong K.K., Kwan H.S.;
 RN Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreill-
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Laym-
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimu-
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -!- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
 (MOLYBDOPTERIN), NECESSARY FOR MOLYBDOENZYMES. PLAYS A ROLE

OF THE SMALL SUBUNIT OF THE MOLYBDOPTEIN CONVERTING
 (AD).
 polypeptide cofactor biosynthesis.
 BELONGS TO THE HES/MOE/THIF FAMILY.
 Entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 res a license agreement (See <http://www.isb-sib.ch/announce/>
 ail to license@isb-sib.ch).

 AAA96530.1; --
 5; AAL19781.1; --
 565; moeb.
 009036; Moeb.
 007901; Moeb Moeb.
 00205; NAD BS.
 00594; Thif domain.
 ; Moeb Moeb; 1.
 ; Thif; 1.
 factor biosynthesis; Complete proteome.
 9 9 M -> I (IN REF. 1).
 38 38 G -> R (IN REF. 1).
 17 117 S -> A (IN REF. 1).
 59 169 N -> T (IN REF. 1).
 21 221 G -> E (IN REF. 1).
 ; AA; 26903 MW; 0F0050831D537AD2 CRC64;
 2.5%; Score 7; DB 1; Length 249;
 arity 100.0%; Pred.No. 60;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 QI 185
 ||||
 QI 16
 STANDARD; PRT; 257 AA.
 rel. 39, Created)
 rel. 39, Last sequence update)
 rel. 41, Last annotation update)
 aride core biosynthesis glycosyl transferase kdtx

 eobacteria; Gammaproteobacteria; Enterobacteriales;
 aceae; Serratia.
 5;
 N.A.
 003; PubMed=8824620;
 pique N.; Climent N., Ferrer S., Merino S., Rubires X.,
 ague M.;
 Characterization of two Serratia marcescens genes
 core lipopolysaccharide biosynthesis.";
 178:5741-5747(1996).
 Lipopolysaccharide core biosynthesis.
 Y: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2. WAAE/KDTX

 Entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 res a license agreement (See <http://www.isb-sib.ch/announce/>
 ail to license@isb-sib.ch).

CC EMBL; US2844; AAC44433.1; --
 DR InterPro; IPR001173; Glyco trans 2.
 DR Pfam; PF00535; Glycos transf 2; 1.
 KW Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase
 SQ SEQUENCE 257 AA; 29233 MW; D40D7B57E002F990 CRC84;
 Query Match 2.5%; Score 7; DB 1; Length 257;
 Best Local Similarity 100.0%; Pred.No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 QY 71 GLLAVV 77
 |||||
 Db 229 GLLAVV 235
 RESULT 47
 CN09 HUMAN STANDARD; PRT; 277 AA.
 AC Q86T03; Q86U09; Q8WUC0; Q9BU67; Q9NSU8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Hypothetical protein Cl4orf9.
 GN Cl4ORF9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Neuroblastoma, and T-cell;
 RA Li W.B., Gruber C., Jessee J., Polayes D.;
 RT "Full-length cDNA libraries and normalization."
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Testis;
 RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
 RA Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain, and Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S
 RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q86T03-1; Sequence=Displayed;
 CC Note=No experimental confirmation available;
 CC Name=2;
 CC IsoId=Q86T03-2; Sequence=VSP_007815;
 CC Note=No experimental confirmation available;
 CC

86T03-3; Sequence-VSP 007816, VSP_007817;
 y be due to intron retention.;

 NOT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 urses a license agreement (See <http://www.isb-sib.ch/announce/>
 ail to license@isb-sib.ch).

 90; CAD61939.1; -;
 95; CAD62347.1; ALT INIT.
 97; CAB70896.1; ALT_INIT.
 97; AAH02867.2; -;
 97; AAH20947.1; -;
 T46382
 .9299; C14orf9.
 protein; Transmembrane; Alternative splicing.
 112 234
 144 266
 47 47
 /FTId-VSP 007815.
 A -> AGKHAPPO (in isoform 2).
 48 224
 APPFPEGHFAVLPGEDPPYSLTSPDSSGAPMITCRVCQ
 SLINVEGKMHQHVKGVCNEATPIKNAPPKKYVRCPCNC
 LLIKVTQRIACPRPKRIINLGFVHEPLSPFPQMGV
 RVICGHCKNTLWTEFTDLARCPCRKRVSSIGRRYPRKR
 CICCFLGLLAV -> GKHPPOGKRGVAGAPRGLTKAG
 EGAPPAEAGPSRQVDCCTCDWARLPSLRNRDHSIGTGG
 SQPDRSANYEPESELQVRVEDQKPPPTTVEHQWCK (in
 isoform 3).
 /FTId-VSP 007816.
 Missing (in isoform 3).
 125 277
 /FTId-VSP 007817.
 P -> T (IN REF. 3; AAH20947).
 17 AA; 29469 MW; A85FE1F736366CBC CRC64;
 .arity 2.5%; Score 7; DB 1; Length 277;
 100.0%; Pred. No. 66;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 LAV 76
 |||||
 LAV 224
 STANDARD; PRT; 310 AA.
 1;
 (Rel. 41, Created)
 (Rel. 41, Last sequence update)
 (Rel. 43, Last annotation update)
 beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
 or ligand superfamily member 3).
 } OR TNFC.
 (Woodchuck).
 ataxia; Chordata; Craniata; Vertebrata; Euteleostomi;
 heria; Rodentia; Sciurognathi; Scuridae; Scurinae;
 995;
 4 N.A.
 1748; PubMed:10721723;
 all E.A., Brown C.L., Cullen J.M.;
 /phorbol-12-myristate-13-acetate (PMA)-induced
 characterization and biological activity.";
 -305(2000).
 : Cytokine that binds to LTBR/TNFRSF3. May play a specific
 immune response regulation. Provides the membrane anchor
 attachment of the heterotrimeric complex to the cell

CC -!- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits
 CC (less prevalent) two LTA and one LTB subunits (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
 CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a coll-
 CC between the Swiss Institute of Bioinformatics and the EMBL out-
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is in
 CC modified and this statement is not removed. Usage by and for
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF096268; AAF34866.1; -;
 CC EMBL; AF095587; AAF34865.1; -;
 CC HSSP; P01374; 1TNR.
 CC InterPro; IPR006053; TNF abc.
 CC InterPro; IPR006052; TNF family.
 CC InterPro; IPR008983; TNF_like.
 CC InterPro; IPR003636; TNF_subf.
 CC Pfam; PF00229; TNF; 1.
 CC PRINTS; PR01234; TNECROSISFCT.
 CC ProDom; PD002012; TNF_subf; 1.
 CC SMART; SM00207; TNF; 1.
 CC PROSITE; PS00251; TNF_1; FALSE_NEG.
 CC PROSITE; PS50049; TNF_2; 1.
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
 FT DOMAIN 1 27
 FT TRANSMEM 28 48
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PRO-
 FT (POTENTIAL).
 FT DOMAIN 49 310
 FT EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 272 272
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 280 280
 FT D -> H (IN REF. 1; AAF34865).
 SQ SEQUENCE 310 AA; 32644 MW; 73B354EFC8B3B3BE CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 QY 194 GLYLYYC 200
 |||||
 DB 181 GLYLYYC 187
 RESULT 49
 ISPH XANCP
 ID ISPH XANCP STANDARD; PRT; 316 AA.
 AC Q8P8G4.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)
 GN ISPH OR LYTB OR XCC1157.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OC NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=20202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.J.
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.I.
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,

Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
Antos M., Truffi D., Tsai S.M., White F.P.,
Kitajima J.P.;
the genomes of two Xanthomonas pathogens with differing
ties";
-463(2002).
Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
ntenyl diphosphate (IPP) and dimethylallyl diphosphate
y similarity).
ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)O =
oxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
onmevalonate terpenoid biosynthesis pathway; seventh
P. Belongs to the ispf family.

T entry is copyright. It is produced through a collabora-
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
rofit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See <http://www.isb-sib.ch/announce/>)
il to license@isb-sib.ch).

; AA040456.1; -
1; -; 1.
03451; LytB.
LYTB; 1.
00216; ispf lytB; 1.
ntthesis; Complete proteome; Oxidoreductase; NADP.
AA; 34641 MW; 6BPSA272A4CA1B4C CRC64;

2.5%; Score 7; DB 1; Length 316;
rity 100.0%; Pred. No. 74;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

LA 223
||
LA 282

STANDARD; PRT; 316 AA.

el. 41, Created)
el. 41, Last sequence update)
el. 42, Last annotation update)
thylbut-2-enyl diphosphate reductase (EC 1.17.1.2).
R XP2416 OR P1435.
iosa, and
iosa (strain Temeculal / ATCC 700964).
eobacteria; Gammaproteobacteria; Xanthomonadales;
ae; Xylella.
1, 183190;
N.A.

17; PubMed:10910347;
Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
onaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
anca S.C., Franco M.C., Frohme M., Furlan L.R.,
ldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
sel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.R.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto R.B., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V. de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
RN [2]

SEQUENCE FROM N.A.
RP STRAIN=Temeculal / ATCC 700964;
RX MEDLINE=22421331; PubMed=12533478;
RC Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.
RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Camargo F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi I
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sassaki F.T., Sena J.A.I
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RL J. Bacteriol. 185:1018-1026(2003).
CC -!- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosph
CC into isopentenyl diphosphate (IPP) and dimethylallyl diphospha
CC (DMAPP) (By similarity).
CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)
CC (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh
CC (last) step.
CC -!- SIMILARITY: Belongs to the ispf family.

This SWISS-PROT entry is copyright. It is produced through a colla
between the Swiss Institute of Bioinformatics and the EMBL outst
the European Bioinformatics Institute. There are no restrictions
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for c
entities requires a license agreement (See <http://www.isb-sib.ch/>
or send an email to license@isb-sib.ch).

EMBL; AE004050; AAF85215.1; -
DR EMBL; AE012558; AAO29279.1; -
DR PIR; C82561; C82561.
DR HAMAP; MF_00191; -; 1.
DR InterPro; IPR003451; LytB.
DR Pfam; PF02401; LYTB; 1.
DR TIGRFAMs; TIGR00216; ispf lytB; 1.
KW Isoprene biosynthesis; Complete proteome; Oxidoreductase; NADP.
SQ SEQUENCE 316 AA; 34704 MW; 1A2E80B9A98D334A CRC64;

Query Match 2.5%; Score 7; DB 1; Length 316;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps
QY 217 LVDGVLA 223
DB 276 LVDGVLA 282
|||||
RESULT 51
ODPB_BACSU STANDARD; PRT; 324 AA.

Rel. 18, Created)
 Rel. 18, Last sequence update)
 Rel. 42, Last annotation update)
 drogenase E1 component, beta subunit (EC 1.2.4.1) (S
 Da subunit)
 OR BSU14590.
 ilis.
 micutes; Bacillales; Bacillaceae; Bacillus.
 23;
 N.A.
 558; PubMed=1697575;
 lva A., Paulin L., Arvidson S., Palva I.;
 complex of *Bacillus subtilis*: sequence analysis and
 yruvate dehydrogenase.";
 172:5052-5063(1990).
 N.A.
 187; PubMed=8969500;
 Aldwell R., Enfield L., Ferrari E.;
 E (124 degrees-127 degrees) region of the *Bacillus*
 chromosome: sequencing of a 27 kb segment and
 n of several genes in the area.";
 142:3033-3037(1996).
 N.A.
 Ferrari E.;
 lysis of the mobA-ampS region of the *Bacillus subtilis*
 IL-1997) to the EMBL/GenBank/DBJ databases.
 N.A.
 033; PubMed=9384377;
 sawara N., Moszer I., Albertini A.M., Alloni G.,
 artero M.G., Bessieres P., Bolotin A., Borchert S.,
 oursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 dani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 levine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 Errington J., Fabret C., Ferrari E., Foulger D.,
 ita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 aser P., Goffeau A., Gollightly E.J., Grandi G.,
 Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 Iolsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 amata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 Koetter P., Koningsstein G., Krogh S., Kumano M.,
 pidus A., Lardinols S., Lauber J., Lazarevic V.,
 line A., Liu H., Masuda S., Mauel C., Medigue C.,
 illado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 il T.M., Portetelle D., Porwollik S., Prescott A.M.,
 Pujic P., Burnelle B., Rapoport G., Rey M., Reynolds S.,
 volta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 ilan E., Schleich S., Schroeder R., Scoffone F.,
 Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 Acconi E., Takagi T., Takahashi H., Takemaru K.,
 Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 chiya S., Vandenberg M., Vannier F., Vassarotti A.,
 abutt R., Wedler E., Wedler H., Weitzenegger T.,
 Vipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 a genome sequence of the Gram-positive bacterium *Bacillus*
 19-256(1997).
 : The pyruvate dehydrogenase complex catalyzes the overall
 n of pyruvate to acetyl-CoA and CO(2). It contains
 copies of three enzymatic components: pyruvate
 anase (E1), dihydrolipoamide acetyltransferase (E2) and

lipoamide dehydrogenase (E3).
 -!- FUNCTION: THE B.SUBTILIS PDH COMPLEX POSSESSES ALSO BRANCHED-
 2-OXOACID DEHYDROGENASE (BCDH) ACTIVITY.
 -!- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-
 acetyl-dihydrolipoamide + CO(2).
 -!- COFACTOR: Thiamine pyrophosphate.
 -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
 This SWISS-PROT entry is copyright. It is produced through a coll
 between the Swiss Institute of Bioinformatics and the EMBL out
 the European Bioinformatics Institute. There are no restriction
 use by non-profit institutions as long as its content is if
 modified and this statement is not removed. Usage by and for
 entities requires a license agreement (See <http://www.isb-sib.ch/>
 or send an email to license@isb-sib.ch).
 EMBL; M57435; AAA62682.1; --.
 EMBL; AF012285; AAC24933.1; --.
 EMBL; Z99111; CAB13332.1; --.
 PIR; C36718; C36718.
 HSSP; P09061; 10S0.
 Subtilist; BG10208; pdhA.
 InterPro; IPR009014; Transketo_C_like.
 InterPro; IPR005476; Transketolase_C.
 InterPro; IPR005475; Transketolase_CR.
 Pfam; PF02779; transket_pyr; 1.
 Pfam; PF02780; transketolase_C; 1.
 Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
 Complete proteome.
 INIT MET 0 BY SIMILARITY.
 SEQUENCE 324 AA; 35343 MW; D2A7C9B32DEDF0D CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 324;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 QY 59 LALGLGL 65
 DB 65 LALGLGL 71
 RESULT 52
 SRA6 CAEBL STANDARD; PRT; 329 AA.
 ID SRA6 CAEBL STANDARD; PRT; 329 AA.
 AC Q09208;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serpentine receptor class alpha 6 (Sra-6 protein).
 GN SRA-6 OR AH6.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoi
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_Taxid=62359;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Jassal B.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SI
 FAMILY.
 This SWISS-PROT entry is copyright. It is produced through a coll
 between the Swiss Institute of Bioinformatics and the EMBL out
 the European Bioinformatics Institute. There are no restriction
 use by non-profit institutions as long as its content is if
 modified and this statement is not removed. Usage by and for
 entities requires a license agreement (See <http://www.isb-sib.ch/>
 or send an email to license@isb-sib.ch).
 EMBL; Z48009; CAA88083.1; --.
 PIR; T18619; T18619.

5. adiodurans.
 nococcus-Thermus; Deinococci; Deinococcales;
 e; Deinococcus.
 99;
 N.A.
 TCC 13939 / DSM 20539 / NCIB 9279;
 896; PubMed=10567266;
 en J.A., Heideberg J.F., Hickey E.K., Peterson J.D.,
 Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 J., Lam P., McDonald L., Utterback T., Zalewski C.,
 Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 nce of the radioresistant bacterium Deinococcus
 1"; 1577(1999).
 Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-
 tanate (3-isopropylmalate) to 3-carboxy-4-methyl-2-
 oate. The product decarboxylates to 4-methyl-2-
 oate.
 ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +
 3-carboxy-4-methyl-2-oxopentanoate + NADH.
 Leucine biosynthesis; third step.
 Homodimer (By similarity).
 AR LOCATION: Cytoplasmic (By similarity).
 Y: Belongs to the isocitrate and isopropylmalate
 nases family. LeuB subfamily 1.
 OT entry is copyright. It is produced through a collaboration
 Swiss Institute of Bioinformatics and the EMBL outstation -
 Bioinformatics Institute. There are no restrictions on its
 profit institutions as long as its content is in no way
 this statement is not removed. Usage by and for commercial
 iles a license agreement (See <http://www.isb-sib.ch/announce/>
 ail to license@isb-sib.ch).
 9; AAF11333.1; -
 G75355.
 1XAA.
 33; -; 1.
 001804; Isodh.
 004429; LeuB.
 ; isodh; 1.
 R00169; leuB; 1.
 470; IDH-IMDH; 1.
 e; Leucine biosynthesis; NAD; Complete proteome.
 2 AA; 37598 MW; 8BAE0E347F2AFA29 CRC64;
 2.5%; Score 7; DB 1; Length 352;
 arity 100.0%; Pred. No. 81;
 onservative 0; Mismatches 0; Indels 0; Gaps 0;
 ALR 254
 |||||
 ALR 94
 STANDARD; PRT; 357 AA.
 Rel. 29, Created)
 Rel. 29, last sequence update)
 Rel. 41, last annotation update)
 sphatase (BC 3.1.3.9) (G6Pase) (G-6-Pase).
 (Mouse).
 tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 heria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94012716; PubMed=8407995;
 RA Shelly L.L., Lei K.-J., Pan C.-J., Sakata S.F., Ruppert S.,
 RA Schutz G., Chou J.Y.;
 RT "Isolation of the gene for murine glucose-6-phosphatase, the enz
 RL deficient in glycogen storage disease type 1A.";
 RN J. Biol. Chem. 268:21482-21485(1993).
 [2]
 RN SEQUENCE OF 1-76 FROM N.A.
 RC STRAIN=129/SV; TISSUE=Liver;
 RX MEDLINE=97277298; PubMed=9115220;
 RA Streeter R.S., Svitek C.A., Chapman S., Greenbaum L.E., Taub R.,
 RA O'Brien R.M.;
 RT "A multicomponent insulin response sequence mediates a strong
 RT repression of mouse glucose-6-phosphatase gene transcription by
 RT insulin.";
 RL J. Biol. Chem. 272:11698-11701(1997).
 CC -!- FUNCTION: May be a single membrane channel protein acting bot
 CC a hydrolase and a translocase. It is the key enzyme in homeos
 CC regulation of blood glucose levels.
 CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate + H(2)O = D-glucose
 CC phosphate.
 CC -!- PATHWAY: Gluconeogenesis and glycogenolysis; last step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum.
 CC -!- TISSUE SPECIFICITY: Liver and kidney.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a coll
 CC between the Swiss Institute of Bioinformatics and the EMBL out
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is ir
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U00445; AAC52122.1; -
 CC EMBL; U91573; AAC53166.1; -
 CC PIR; A48589; A48589.
 CC MGD; MGI:95607; G6pc.
 CC InterPro; IPR008934; AcPase_VanPerase.
 CC InterPro; IPR000326; PA_PTPase.
 CC Pfam; PF01569; PAP2; 1.
 CC SMART; SM00014; acidPPc; 1.
 CC Glycogen biosynthesis; Hydrolase; Transmembrane; Glycoprotein;
 CC Endoplasmic reticulum.
 CC TRANSMEM 30 46 POTENTIAL.
 CC TRANSMEM 59 75 POTENTIAL.
 CC TRANSMEM 153 169 POTENTIAL.
 CC TRANSMEM 211 227 POTENTIAL.
 CC TRANSMEM 296 312 POTENTIAL.
 CC TRANSMEM 333 349 POTENTIAL.
 CC TRANSMEM 96 96 N-LINKED (GLCNAC. . .) (BY SIMILARIT
 CC CARBOHYD 354 357 PREVENT SECRETION FROM ER (POTENTIAL
 CC SITE 354 357
 CC SEQUENCE 357 AA; 40480 MW; 292F9FCE39582692 CRC64;
 SQ
 Query Match 2.5%; Score 7; DB 1; Length 357;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 QY 61 LGGLAL 67
 Db 269 LGGLAL 275
 |||||
 |||||
 RESULT 57
 G6PT RAT STANDARD; PRT; 357 AA.
 ID_G6PT RAT
 AC P43428;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

el. 41, Last annotation update)
phatase (EC 3.1.3.9) (G6Pase) (G-6-Pase).

cus (Rat).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
16;

N.A.

19; PubMed=7860767;
in S., Chung E., Buikuisen W., Naji A., Taub R.A.;
f glucose-6-phosphatase gene and protein expression
tive response in proliferating liver and diabetes.";
t. 95:832-841(1995).

N.A.

-Dawley;
95; PubMed=8198588;
gaud D.M., El-Maghrabi M.R., Pan W., Subir M.,

a cDNA for the catalytic subunit of rat liver
phatase: regulation of gene expression in FAO hepatoma
in, dexamethasone and cAMP.";
ys. Res. Commun. 201:302-309(1994).

N.A.

-Dawley; TISSUE=Liver;
50; PubMed=8865366;
ajima H., Horikawa Y., Hamaguchi T., Yamasaki T.,
Iamba M., Hanafusa T., Matsuzawa Y.;
d distribution of glucose-6-phosphatase catalytic
ger RNA and its changes in the diabetic state.";
ol. Pathol. Pharmacol. 93:13-24(1996).
May be a single membrane channel protein acting both as
e and a translocase. It is the key enzyme in homeostatic
of blood glucose levels.
ACTIVITY: D-glucose 6-phosphate + H(2)O = D-glucose +

luconeogenesis and glycogenolysis; last step.

R LOCATION: Integral membrane protein. Endoplasmic

T entry is copyright. It is produced through a collaboration
wiss Institute of Bioinformatics and the EMBL outstation -
ioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (See <http://www.isb-sib.ch/announce/>
il to license@isb-sib.ch).

AA474381.1; ALT_INIT.

AA419966.1; -.

BAA24348.1; -.

C2371.

08934; AcPase VanPerase.

00326; PA_PTPase.

PAP2; l.

; acidppc; l.

ntesis; Hydrolase; Transmembrane; Glycoprotein;

ticulum.

0 46 POTENTIAL.

9 75 POTENTIAL.

3 169 POTENTIAL.

1 227 POTENTIAL.

6 312 POTENTIAL.

3 349 POTENTIAL.

6 96 N-LINKED (GLCNAC. . .) (BY SIMILARITY).

4 357 PREVENT SECRETION FROM ER (POTENTIAL).

8 118 G -> V (IN REF. 2).

AA; 40555 MW; C44960E102F4244D CRC64;

2.5%; Score 7; DB 1; Length 357;

Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; G;

QY 61 LGLGLAL 67

Db 269 LGLGLAL 275

RESULT 58

PONI_RABIT

ID PONI_RABIT STANDARD; PRT; 358 AA.

AC P27170; Q9BGN1; Q9BGN2; Q9BGN3;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Serum paraoxonase/arylesterase 1 (EC 3.1.1.2) (EC 3.1.8.1) (PONI 1;

DE (Serum arylalkylphosphatase 1) (A-esterase 1) (Aromatic esterase

GN PONI OR PON.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Liver;

RX MEDLINE=92031445; PubMed=1657140;

RA Hassett C., Richter R.J., Humbert R., Chapline C., Crabb J.W.,

RA Omiecinski C.J., Furlong C.E.;

RT "Characterization of cDNA clones encoding rabbit and human serum

RT paraoxonase: the mature protein retains its signal sequence.";

RL Biochemistry 30:10141-10149(1991).

RN [2]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RX MEDLINE=93345100; PubMed=8393745;

RA Furlong C.E., Costa L.G., Hassett C., Richter R.J.,

RA Sundstrom J.A., Adler D.A., Distche C.M., Omiecinski C.J.,

RA Chapline C., Crabb J.W.;

RT "Human and rabbit paraoxonases: purification, cloning, sequencing

RT mapping and role of polymorphism in organophosphate detoxification

RL Chem. Biol. Interact. 87:35-48(1993).

RN [3]

RP SEQUENCE FROM N.A., FUNCTION, AND VARIANTS SER-81; GLU-92 AND GLY.

RC STRAIN=New Zealand white; TISSUE=Liver;

RX MEDLINE=21163843; PubMed=11266077;

RA Watson C.E., Braganov D.I., Billecke S.S., Bisgaier C.L., La Du B

RT "Rabbits possess a serum paraoxonase polymorphism similar to the 1

RT Q192R.";

RL Pharmacogenetics 11:123-134(2001).

RN [4]

RP CHARACTERIZATION, AND SEQUENCE OF 1-20.

RX MEDLINE=92031444; PubMed=1718413;

RA Furlong C.E., Richter R.J., Chapline C., Crabb J.W.;

RT "Purification of rabbit and human serum paraoxonase.";

RL Biochemistry 30:10133-10140(1991).

CC -!- FUNCTION: Hydrolyzes the toxic metabolites of a variety of

CC organophosphorus insecticides. Capable of hydrolyzing a broad

CC spectrum of organophosphate substrates and a number of aromatic

CC carboxylic acid esters. Mediates an enzymatic protection of l

CC density lipoproteins against oxidative modification.

CC -!- CATALYTIC ACTIVITY: Aryl dialkyl phosphate + H(2)O = dialkyl

CC phosphate + an aryl alcohol.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- PTM: Glycosylated.

CC -!- PTM: The signal sequence is not cleaved.

CC -!- POLYMORPHISM: There are two allelic forms, allozyme A and B, v

CC differ in their substrate specificity. Both forms have similar;

CC arylesterase activity but allozyme B possesses greater paraox

CC activity. Allozyme A is better at protecting LDL from oxidati

CC -!- SIMILARITY: Belongs to the paraoxonase family.

CC -!- CAUTION: Ref.3 (AAK06398) sequence differs from that shown due

CC a stop codon in position 355.

OT entry is copyright. It is produced through a collaboration Swiss Institute of Bioinformatics and the EMBL outstation - Bioinformatics Institute. There are no restrictions on its profit institutions as long as its content is in no way this statement is not removed. Usage by and for commercial purposes is not removed. (See <http://www.isb-sib.ch/announce/> for license agreement)

AA31452.1; -
AA27713.2; -
1; AAK06398.1; ALT_TERM.
2; AAK06399.1; -
3; AAK06400.1; -
B40354.
6; C:extracellular; NAS.
9; F:antioxidant activity; IDA.
3; F:arylaliphosphatase activity; NAS.
4; F:arylesterase activity; IDA.
3; F:response to organophosphorus; IDA.
002640; Arylesterase.
008363; Paraoxonase1.
008364; Paraoxonase2.
; Arylesterase; 1.
85; PARAOXONASE.
86; PARAOXONASE1.
87; PARAOXONASE2.
tioxidant; Glycoprotein; Signal; Multigene family;

0 0 ?
1 352
41 BY SIMILARITY.
49 N-LINKED (GLCNAC. .) (POTENTIAL).
52 N-LINKED (GLCNAC. .) (POTENTIAL).
252 N-LINKED (GLCNAC. .) (POTENTIAL).
69 N-LINKED (GLCNAC. .) (POTENTIAL).
23 N-LINKED (GLCNAC. .) (POTENTIAL).
81 P -> S (IN ALLELE A).
92 K -> E (IN ALLELE A).
100 S -> G (IN ALLELE A).
66 A -> V (IN REF. 3).
19 319 A -> V (IN REF. 3); AAK06398).
8 AA; 39878 MW; C40C46S5F6S5E5FDF CRC64;

2.5%; Score 7; DB 1; Length 358;
arity 100.0%; Pred.No. 82;
conservative 0; Mismatches 0; Indels 0; Gaps 0;

LAL 67
LAL 15

STANDARD; PRT; 365 AA.

Rel. 35, Created)
Rel. 35, Last sequence update)
Rel. 43, Last annotation update)
oxidoreductase chain 8 (EC 1.6.99.5) (NADH dehydrogenase
NDH-1, chain 8).

ophilus.
nococcus-Thermus; Deinococci; Thermales; Thermaceae;
4;

1 N.A.
ATCC 27634;
490; PubMed=9020134;
S.S., Sled, V.D., Ohnishi T., Yagi T.;
translocating NADH-quinone oxidoreductase (NDH-1) of
bacterium Thermus thermophilus HB-8. Complete DNA

RT sequence of the gene cluster and thermostable properties of the
expressed NQO2 subunit."
J. Biol. Chem. 272:4201-4211 (1997).
-!- FUNCTION: NQO-1 shuttles electrons from NADH, via FMN and iron
sulfur (Fe-S) centers, to quinones in the respiratory chain.
Immediate electron acceptor for the enzyme in this species is
believed to be menaquinone. Couples the redox reaction to proton
translocation (for every two electrons transferred, four protons
are translocated across the cytoplasmic membrane), and thus
conserves the redox energy in a proton gradient.
-!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
-!- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS. SUBUNIT NQO7-14
CONSTITUTE THE MEMBRANE SECTOR OF THE COMPLEX.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: Belongs to the complex I subunit 1 family.

This SWISS-PROT entry is copyright. It is produced through a coll
between the Swiss Institute of Bioinformatics and the EMBL out
the European Bioinformatics Institute. There are no restrictions
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for c
entities requires a license agreement (See <http://www.isb-sib.ch/>
or send an email to license@isb-sib.ch).

EMBL; U52917; AAA97945.1; -
PIR; T11905; T11905.
DR InterPro; IPR001694; Resp_NADH_dh1.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_NDI_1; 1.
DR PROSITE; PS00668; COMPLEX1_NDI_2; 1.
KW Oxidoreductase; NAD; Quinone; Transmembrane.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 252 272 POTENTIAL.
FT TRANSMEM 273 293 POTENTIAL.
FT TRANSMEM 310 330 POTENTIAL.
FT TRANSMEM 336 356 POTENTIAL.
SQ SEQUENCE 365 AA; 41008 MW; AE920CC029333C09 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 365;
Best Local Similarity 100.0%; Pred.No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 61 LGLGLAL 67
Db 164 LGLGLAL 170

RESULT 60

BENE ACICA
ID BENE ACICA STANDARD; PRT; 394 AA.
AC P07775;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Benzoate membrane transport protein.
GN BENE.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BD413 / ADP1;
RX MEDLINE=91358314; PubMed=188518;
RA Neidle E.L., Hartnett C., Ornston N.L., Bairoch A., Rekik M.,
RA Harayama S.;
FT "Nucleotide sequences of the Acinetobacter calcoaceticus benABC
RT for benzoate 1,2-dioxygenase reveal evolutionary relationships an
multicomponent oxygenases.";

852; PubMed=7671800;
McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
achy P.A.;
pression and shared activities of members of the hedgehog
of *Xenopus laevis*.;
21:2337-2347(1995).
SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF
SAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC
AND FORMATION IN EMBRYOS.
AR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
LE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
FACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED
VED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY
Y).
C-terminal domain displays an autophosphorylation activity
lesterol transferase activity. Both activities result in
age of the full-length protein and covalent attachment of
erol moiety to the C-terminal of the newly generated N-
fragment (N-product). This covalent modification appears
in essential role in restricting the spatial distribution
rotein activity to the cell surface. The N-product is the
pecies in both local and long-range signaling, whereas the
has no signaling activity (By similarity).
Y: Belongs to the hedgehog family.

OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ures a license agreement (See <http://www.isb-sib.ch/announcement/>
mail to license@isb-sib.ch).

AA085164.1; -
1VHH.
IPW; -
R009045; Hedgehog/DD_pept.
R003587; Hedgehog hint N.
R003586; Hedgehog hintC.
R000320; HH signal.
R001767; Pept C46 hint.
R001657; Peptidase_C46.
; HH signal; 1.
; Hint; 1.
32; SONICHHOG.
042; HH signal; 1.
5; HintG; 1.
6; HintN; 1.
- protein; Autocatalytic cleavage; Hydrolase; Protease;
- protein; Palmitate.
1 23 POTENTIAL.
24 398 DESERT HEDGEHOG PROTEIN 2.
24 199 DESERT HEDGEHOG PROTEIN 2 N-PRODUCT.
200 398 DESERT HEDGEHOG PROTEIN 2 C-PRODUCT.
278 281 POLY-SER.
199 200 CLEAVAGE (AUTO-) (BY SIMILARITY).
269 269 INVOLVED IN AUTO-CLEAVAGE (BY
SIMILARITY).
272 272 ESSENTIAL FOR AUTO-CLEAVAGE (BY
SIMILARITY).
24 24 N-palmitoyl cysteine (By similarity).
199 199 Cholesterol glycine ester (By
similarity).
98 AA; 44458 MW; DBC23AF85F69DD08 CRC64;
2.5%; Score 7; DB 1; Length 398;
larity 100.0%; Pred.No.90;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CGVL 222
|||
CGVL 339

RESULT 63
SELP_BOVIN
ID SELP_BOVIN STANDARD; PRT; 402 AA.
AC P49907; O19003;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Selenoprotein P-like protein precursor.
GN SEPL OR SELP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=95364621; PubMed=7637580;
RA Saijoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.
RT "Molecular cloning of cDNA encoding a bovine selenoprotein P-like
protein containing 12 selenocysteines and a (His-Pro) rich domain
insertion, and its regional expression.";
RT Brain Res. Mol. Brain Res. 30:301-311(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98019090; PubMed=9358058;
RA Fujii M., Saijoh K., Kobayashi T., Fujii S., Lee M.J., Sumino K.
RT "Analysis of bovine selenoprotein P-like protein gene and availa
of metal responsive element (MRE) located in its promoter.";
RL Gene 199:211-217(1997).
CC -!- FUNCTION: It constitutes a major selenium pool in the brain;
may play an important role in developing and/or modulating t
morphology of neurons and/or glial cells.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: Brain and kidney. Most prominently expre
in the cerebellar cortex, hippocampus and olfactory bulb.
CC -!- MISCELLANEOUS: The selenocysteines are all encoded by the op
codon, UGA.

This SWISS-PROT entry is copyright. It is produced through a col
between the Swiss Institute of Bioinformatics and the EMBL out
the European Bioinformatics Institute. There are no restriction
use by non-profit institutions as long as its content is in
modified and this statement is not removed. Usage by and for
entities requires a license agreement (See <http://www.isb-sib.ch>
or send an email to license@isb-sib.ch).

EMBL; D25220; BAA04949.2; -
EMBL; D88033; BAA23414.1; -
EMBL; D88031; BAA23414.1; JOINED.
EMBL; D88032; BAA23414.1; JOINED.
InterPro; IPR007672; Selp_C.
InterPro; IPR007671; Selp_N.
Pfam; PF04593; Selp_C; 1.
Pfam; PF04592; Selp_N; 1.
KW Glycoprotein; Signal; Selenium; Selenocysteine; Repeat.
SIGNA 1 19
FT CHAIN 20 402
FT SE_CYS 59 59
FT SE_CYS 297 297
FT SE_CYS 307 307
FT SE_CYS 338 338
FT SE_CYS 350 350
FT SE_CYS 363 363
FT SE_CYS 365 365
FT SE_CYS 372 372
FT SE_CYS 388 388
FT SE_CYS 390 390
FT SE_CYS 397 397
FT SE_CYS 399 399
FT DOMAIN 204 239
H-P REPEATS.

```

0 266 POLY-HIS.
8 181 SRPQ -> KALE (IN REF. 2).
6 256 T -> P (IN REF. 2).
2 282 L -> V (IN REF. 2).
2 312 Y -> D (IN REF. 2).
AA: 45018 MW; B7CF18751F808BFF CRC64;
      2.5%; Score 7; DB 1; Length 402;
      100.0%; Pred. No. 91;
      0; Mismatches 0; Indels 0; Gaps 0;
IA 68
II
IA 10

      STANDARD; PRT; 412 AA.

el. 34, Created)
el. 34, Last sequence update)
el. 41, Last annotation update)
te kinase, plasmid (EC 2.7.2.3).

trophus (Ralstonia eutropha).
asid pHG1.
eobacteria; Betaproteobacteria; Burkholderiales;
ae; Ralstonia.

N.A.
ISM 428 / ATCC 17699;
15; PubMed=7763137;
1 J., Yoo J.-G., Bowien B.;
the genes forming the distal parts of the two cbb CO2
1. 163:291-299(1995).
Activity: ATP + 3-phospho-D-glycerate = ADP + 3-
glyceroyl phosphate.
Calvin cycle.
monomer (By similarity).
R LOCATION: Cytoplasmic.
: Belongs to the phosphoglycerate kinase family.

Entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstati-
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
his statement is not removed. Usage by and for commercial
res a license agreement (see http://www.isb-sib.ch/announce/
il to license@isb-sib.ch).

AAC43447.1; -.
1PHP.
15; -. 1.
001576; PGK.
PGK; 1.
7; PGLYCKINASE.
11; PGLYCERATE_KINASE; 1.
Kinase; Calvin cycle; Plasmid.
2 AA; 42298 MW; 958406685957274 CRC64;
      2.5%; Score 7; DB 1; Length 412;
      100.0%; Pred. No. 93;
      0; Mismatches 0; Indels 0; Gaps 0;
TAA 235
III
TAA 293

```

```

PGKC_ALCEU
ID PGKC_ALCEU STANDARD; PRT; 413 AA.
AC P50319;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase, chromosomal (EC 2.7.2.3).
GN CBBKC.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=510;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=H16 / DSM 428 / ATCC 17699;
RX MEDLINE=95283415; PubMed=7763137;
RA Schaeferfohann J., Yoo J.-G., Bowien B.;
RT Analysis of the genes forming the distal parts of the two cbb CO2
fixation operons from Alcaligenes eutrophus.";
RL Arch. Microbiol. 163:291-299(1995).
CC -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
phospho-D-glyceroyl phosphate.
CC -!- PATHWAY: Calvin cycle.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.

This SWISS-PROT entry is copyright. It is produced through a colla-
between the Swiss Institute of Bioinformatics and the EMBL outstati-
the European Bioinformatics Institute. There are no restrictions on
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).

EMBL; U12422; AAC43444.1; -.
DR PIR; I39551; I39551.
DR HSSP; P18912; 1PHP.
DR HAMAP; MF_00145; -. 1.
DR InterPro; IPR001576; PGK.
DR Pfam; PF00162; PGK; 1.
DR PRINTS; PR00477; PGLYCKINASE.
DR PROSITE; PS00111; PGLYCERATE_KINASE; 1.
KW Transferase; Kinase; Calvin cycle.
SQ SEQUENCE 413 AA; 42283 MW; 6B4C9D195566A90D CRC64;
      2.5%; Score 7; DB 1; Length 413;
      100.0%; Pred. No. 93;
      0; Mismatches 0; Indels 0; Gaps 0;
QY 229 EFSATAA 235
      |||||
DB 288 EFSATAA 294

RESULT 66
NH59 CAEEL STANDARD; PRT; 416 AA.
AC QSTXU1; Q9GTF2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nuclear hormone receptor family member nhr-59.
GN NHR-59 OR T27B7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoid;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RP Bogan A., Maina C.V., Yamamoto K., Cohen P., Sluder A.E.;
RA "Caenorhabditis elegans nuclear receptor sequences exhibit biophy-
RT compatibility with the ligand-binding domain fold.";

```


rophilum.
 rchaeta; Thermoprotei; Thermoproteales;
 ae; Pyrobaculum.
 73;
 N.A.
 TCC 51768 / DSM 7523;
 97; PubMed=11792869;
 T.; Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 ce of the hyperthermophilic crenarchaeon Pyrobaculum
 ad. Sci. U.S.A. 99:984-989 (2002).
 ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
 Magnesium is required for catalysis and for stabilizing
 (By similarity).
 Glycolysis.
 homodimer (By similarity).
 R LOCATION: Cytoplasmic (By similarity).
 : Belongs to the enolase family.
 T entry is copyright. It is produced through a collabora-
 tion between the Swiss Institute of Bioinformatics and the EMBL out-
 standing Bioinformatics Institute. There are no restrictions
 on its use by non-profit institutions as long as its content is in
 profit institutions as long as its content is in no way
 his statement is not removed. Usage by and for commercial
 res a license agreement (See <http://www.isb-sib.ch/announce/>
 ul to license@isb-sib.ch).
 : AAL63046.1; -.
 8; -; 1.
 00941; Enolase.
 enolase; 1.
 enolase N; 1.
 8; ENOLASE.
 02; Enolase; 1.
 64; ENOLASE; FALSE_NEG.
 'ais; Magnesium; Complete proteome.
 BY SIMILARITY.
 1 151
 MAGNESIUM (BY SIMILARITY).
 0 240
 MAGNESIUM (BY SIMILARITY).
 3 283
 MAGNESIUM (BY SIMILARITY).
 9 309
 MAGNESIUM (BY SIMILARITY).
 1 AA; 43377 MW; 186068A3137D7F9E CRC64;
 2.5%; Score 7; DB 1; Length 419;
 ity 100.0%; Pred. No. 95;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 P 16
 ||
 P 141
 STANDARD; PRT; 419 AA.
 el. 41, Created)
 el. 41, Last sequence update)
 el. 41, Last annotation update)
 re kinase (EC 2.7.2.3).
 l OR RS04894.
 inacearum (Pseudomonas solanacearum).
 eobacteria; Betaproteobacteria; Burkholderiales;
 ae; Ralstonia.
 5;
 N.A.
);
 379; PubMed=11823852;
 . Genin S., Artiguenave F., Gouzy J., Mangenot S.,

RA Arlat M., Billault A., Brottier P., Camus J.-C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502 (2002).
 CC -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
 CC phospho-D-glyceroyl phosphate.
 CC -!- PATHWAY: Second phase of glycolysis; second step.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
 CC This SWISS-PROT entry is copyright. It is produced through a colla-
 CC between the Swiss Institute of Bioinformatics and the EMBL out-
 CC the European Bioinformatics Institute. There are no restrictions
 CC use by non-profit institutions as long as its content is in
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AL646060; CAD14101.1; -.
 DR HAMAP; MF_00145; -; 1.
 DR InterPro; IPR001576; PGK.
 DR Pfam; PF00162; PGK; 1.
 DR PRINTS; PR00477; PHGLYCKINASE.
 DR PROSITE; PS00111; PGLYCERATE_KINASE; 1.
 DR Transferase; Kinase; Glycolysis; Complete proteome.
 KW SEQUENCE 419 AA; 43268 MW; 8219241085E3D81 CRC64;
 SQ
 Query Match 2.5%; Score 7; DB 1; Length 419;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; G
 QY 229 EFSATAA 235
 |||||
 Db 289 EFSATAA 295
 RESULT 70
 YJ54 YEAST
 ID YJ54 YEAST STANDARD; PRT; 423 AA.
 AC P47130;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Hypothetical 49.5 kDa protein in MIR1-STE18 intergenic region.
 GN YJ084W OR J1860.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=96437976; PubMed=8840504;
 RA Huang M.-E., Manus V., Chuat J.-C., Galibert F.;
 RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
 RT reading frames and a gene cluster with a counterpart on chromosom
 RT XI.";
 RL Yeast 12:869-875 (1996).
 CC This SWISS-PROT entry is copyright. It is produced through a colla-
 CC between the Swiss Institute of Bioinformatics and the EMBL out-
 CC the European Bioinformatics Institute. There are no restrictions
 CC use by non-profit institutions as long as its content is in
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z49584; CAA89611.1; -.
 DR EMBL; L47993; AAB39307.1; -.

SS7103.
 41916; --
 ; YJUR084W.
 0; C:signalosome complex; IDA.
 4; P:adaptation to pheromone during conjugation . . .; IMP.
 000717; PCI.
 ; PCI; 1.
 protein.
 3 AA; 49482 MW; 750CDA631916A621 CRC64;
 2.5%; Score 7; DB 1; Length 423;
 arity 100.0%; Pred. No. 95;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 LAL 61
 ||||
 LAL 272
 STANDARD; PRT; 424 AA.
 Rel. 11, Created
 Rel. 35, Last sequence update
 Rel. 39, Last annotation update
 a sperm-binding protein 3 precursor (Zona pellucida
 ZP3) (sperm receptor) (Zona pellucida protein C).
 (Mouse).
 tazia; Chordata; Craniata; Vertebrata; Euteleostomi;
 heria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 090;
 (N.A.
 926; PubMed=3378665;
 ; Chamberlin M.E.; Baur A.W.; Sobieski D.A.; Dean J.;
 analysis of cDNA coding for ZP3, a sperm binding protein
 zona pellucida.";
 7:287-295 (1988).
 87.
 V-1996) to the EMBL/GenBank/DBJ databases.
 (N.A.
 TISSUE=Liver;
 048; PubMed=2541416;
 Wassarman P.M.;
 sequence of the gene encoding zona pellucida glycoprotein
 use sperm receptor.";
 Res. 17:2861-2863 (1989).
 (N.A.
 451; PubMed=2842770;
 Roller R.J.; Fimiani C.M.; Wassarman D.A.;
 ;
 cture of the mouse sperm receptor polypeptide determined
 oning.";
 acad. Sci. U.S.A. 85:6409-6413 (1988).
 19-63; 197-204; 219-233 AND 261-275.
 19795; PubMed=1330788;
 Wassarman P.M.;
 on of a region of mouse zona pellucida glycoprotein mzp3
 is sperm receptor activity.";
 14:309-317 (1992).
 ; Functions as a sperm-receptor. It is responsible for
 sion to the zona pellucida, and may contribute to the
 specificity of the insemination.
 ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
 ; AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
 CC matrix.
 CC -!- TISSUE SPECIFICITY: Oocytes.
 CC -!- DEVELOPMENTAL STAGE: Expressed during the 2-week growth phase
 CC oogenesis, prior to ovulation.
 CC -!- PM: Sulfated glycoprotein with O-linked oligosaccharides.
 CC -!- SIMILARITY: Contains 1 ZP domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a coll
 CC between the Swiss Institute of Bioinformatics and the EMBL out
 CC the European Bioinformatics Institute. There are no restriction
 CC use by non-profit institutions as long as its content is ir
 CC modified and this statement is not removed. Usage by and for c
 CC entities requires a license agreement (See <http://www.isb-sib.ch/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M20026; AAB18629.1; --
 CC EMBL; X14376; CAA32550.1; --
 CC PIR; A30334; A30334.
 CC MGD; MGI:99215; Zp3.
 CC InterPro; IPR001507; Endoglin/CD105.
 CC Pfam; PF00100; zona_pellucida; 1.
 CC PRINTS; PR00023; ZPELLUCIDA.
 CC SMART; SM00241; ZP; 1; DOMAIN; 1.
 CC PROSITE; PS00682; ZP_DOMAIN; 1.
 CC Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
 KW Extracellular matrix.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN
 FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 388 408 POTENTIAL.
 FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 45 308 ZP.
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 424 AA; 46303 MW; 9089903FBD268365 CRC64;
 Query Match 2.5%; Score 7; DB 1; Length 424;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; C
 QY 60 ALGLGLA 66
 Db 389 ALGLGLA 395
 |||||
 |||||
 RESULT 72
 SYH CHLMU
 ID SYH CHLMU STANDARD; PRT; 428 AA.
 AC Q9PUJ9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
 DE (Hiers).
 GN HISS OR TC0830.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MoPn / Nigg;
 EX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg
 RA Eisen J., Fraser C.M.;
 RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RA pneumoniae AR39.";

AR LOCATION: Integral membrane protein.
OT entry is copyright. It is produced through a collaboration
Swiss Institute of Bioinformatics and the EMBL outstation -
Bioinformatics Institute. There are no restrictions on its
profit institutions as long as its content is in no way
this statement is not removed. Usage by and for commercial
ures a license agreement (See <http://www.isb-sib.ch/announce/>
ail to license@isb-sib.ch).

AA26163.1; ALT_INIT.

C33958. PUCC.

904896; PUCC.

ex; Transmembrane.

1 36 CYTOPLASMIC (PROBABLE).

37 56 PROBABLE.

57 62 PERIPLASMIC (PROBABLE).

63 83 PROBABLE.

84 109 CYTOPLASMIC (PROBABLE).

10 129 PROBABLE.

30 142 PERIPLASMIC (PROBABLE).

43 153 PROBABLE.

64 182 CYTOPLASMIC (PROBABLE).

83 202 PROBABLE.

103 209 PERIPLASMIC (PROBABLE).

110 228 PROBABLE.

129 261 CYTOPLASMIC (PROBABLE).

162 281 PROBABLE.

182 302 PERIPLASMIC (PROBABLE).

193 319 PROBABLE.

120 338 CYTOPLASMIC (PROBABLE).

139 355 PROBABLE.

156 358 PERIPLASMIC (PROBABLE).

159 376 PROBABLE.

177 394 CYTOPLASMIC (PROBABLE).

195 415 PROBABLE.

116 436 PERIPLASMIC (PROBABLE).

137 456 PROBABLE.

157 461 CYTOPLASMIC (PROBABLE).

11 AA; 48392 MW; BFC7A8A0C549875A CRC64;

2.5%; Score 7; DB 1; Length 461;

arity 100.0%; Pred. No. 1e+02;

conservative 0; Mismatches 0; Indels 0; Gaps 0;

EDQ 101

EDQ 176

STANDARD; PRT; 461 AA.

(Rel. 35, Created)

(Rel. 35, Last sequence update)

(Rel. 41, Last annotation update)

protein HI0608.

influenzae.

teobacteria; Gammaproteobacteria; Pasteurellales;

ae; Haemophilus.

27;

4 N.A.

W20 / ATCC 51907;

3630; PubMed=7542800;

z.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

z., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

, Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus infl

RD.":

RL Science 269:496-512(1995).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- SIMILARITY: BELONGS TO THE SLC13A FAMILY OF TRANSPORTERS.

CC NADC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a coll

CC between the Swiss Institute of Bioinformatics and the EMBL out

CC the European Bioinformatics Institute. There are no restriction

CC use by non-profit institutions as long as its content is li

CC modified and this statement is not removed. Usage by and for

CC entities requires a license agreement (See <http://www.isb-sib.ch/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U32743; AAC22267.1; -.

DR PIR; I64080; I64080.

DR TIGR; HI0608; -.

DR InterPro: IPR001898; Na/sul_sympo.

DR Pfam: PF00939; Na sulph_symp; 1.

DR TIGRFAMs; TIGR00785; class; 1.

DR PROSITE; PS01271; NA_SULFATE; 1.

KW Hypothetical protein; Transmembrane; Transport; Complete proteom

FT TRANSMEM 13 33 POTENTIAL.

FT TRANSMEM 54 74 POTENTIAL.

FT TRANSMEM 81 101 POTENTIAL.

FT TRANSMEM 120 140 POTENTIAL.

FT TRANSMEM 170 190 POTENTIAL.

FT TRANSMEM 211 231 POTENTIAL.

FT TRANSMEM 256 276 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT TRANSMEM 314 334 POTENTIAL.

FT TRANSMEM 349 369 POTENTIAL.

FT TRANSMEM 377 397 POTENTIAL.

FT TRANSMEM 399 419 POTENTIAL.

FT TRANSMEM 439 459 POTENTIAL.

SQ SEQUENCE 461 AA; 49761 MW; B5B6F6965B38EF06 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 461;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; (

QY 53 TALLVPL 59

Db 60 TALLVPL 66

Search completed: April 7, 2004, 17:57:56

Job time : 21 secs

16:25:21 2004

ua-09-245-198a-4.oligo.rspt

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

n search, using sw model

il 7, 2004, 17:54:13 ; Search time 45 Seconds
(without alignments)
1991.270 Million cell updates/sec

09-245-198A-4

MSLDFEISARRLPRLSLG.....PWAHLKAAPLTYFGLFQVH 284

GO

xop 60.0 , Gapext 60.0

.7041 seqs, 315518202 residues

s satisfying chosen parameters: 1017041

jth: 0

jth: 2000000000

.sting first 100 summaries

TREMBL 25:*

sp archaea:*

sp bacteria:*

sp fungi:*

sp human:*

sp invertebrate:*

sp mammal:*

sp_mhc:*

sp_organelle:*

sp_phage:*

sp_plant:*

sp rodent:*

sp virus:*

sp vertebrate:*

sp_unclassified:*

sp_rvirus:*

sp_bacteriap:*

sp_archaeap:*

the number of results predicted by chance to have a
r than or equal to the score of the result being printed,
ad by analysis of the total score distribution.

SUMMARIES

ary	ch	Length	DB	ID	Description
3.5	330	4	Q81ZK7		Q81ZK7 homo sapien
1.3	410	11	Q8BXS2		Q8BXS2 mus musculus
4.2	438	16	Q7VVB7		Q7VVB7 bordetella
4.2	470	16	Q7W7P2		Q7W7P2 bordetella
3.5	111	16	Q8XAT8		Q8XAT8 escherichia
3.5	111	16	Q8PFL8		Q8PFL8 escherichia
3.5	111	16	Q7UC61		Q7UC61 shigella fl
3.2	142	16	Q981J0		Q981J0 rhizobium l
3.2	749	16	Q7V511		Q7V511 prochloroc
3.2	766	16	Q8PZ28		Q8PZ28 xanthomonas
3.2	1208	16	Q7ULK4		Q7ULK4 rhodospirell
2.8	143	17	Q9HST7		Q9HST7 halobacteri
2.8	151	10	Q9SD11		Q9SD11 oryza sativ
2.8	154	2	Q848K4		Q848K4 gamma-prote
2.8	190	16	Q8E569		Q8E569 streptococc
2.8	190	16	Q8DZK5		Q8DZK5 streptococc

Q8L4K2 ory
Q9WYU0 the
Q9VW70 dros
Q8SXH4 dros
Q8fV59 bru
Q89Gw9 bra
Q84MB7 ara
Q9ZG99 pse
Q82143 str
Q9D378 mus
Q9CPR8 mus
Q9VNP0 dros
Q864L1 pan
Q864L0 pong
Q864J8 maca
Q864J7 maca
Q864J5 maca
Q864G8 actu
Q864G7 atel
Q864G6 alou
Q864G4 alou
Q864G3 alou
Q864G2 alou
Q864G1 alou
Q864F8 vare
Q864F7 vare
Q864F6 eule
Q864F5 hapa
Q864F4 lemu
Q85554 pyr
Q9908 homo
Q9BPV2 homo
Q8TUU8 met
Q53860 myc
Q7U145 myc
Q877C5 meth
Q9KYT3 str
Q9SHD8 ara
Q8U820 agr
Q7U9H4 syn
Q8X570 ral
Q8ZCV8 yer
Q9R18 dei
Q96N66 homo
Q8R1P9 mus
Q9CY76 mus
Q8CHK3 mus
Q8GVZ8 ory
Q7WZ71 nono
Q8PMH8 xan
Q9YTU9 inf
Q8JN92 inf
Q8QPL0 inf
Q8PAS2 xan
Q9N8H2 tryp
Q9N8U8 tryp
Q9BEF6 capr
Q9ZG35 chla
Q16193 homo
Q7U6G4 syn
Q8RQ54 serr
Q8H142 ara
Q9A4G4 cau
Q9P074 homo
Q82CB6 nit
Q7X682 ory
P97199 esch
Q8WQ88 eupr
Q8QRH6 hep
Q8UTL1 hum
Q8DBE3 vib
Q38214 bact

2.5 115 16 Q9HY60
 2.5 118 2 Q939F7
 2.5 118 16 Q928C5
 2.5 119 10 Q94LX6
 2.5 121 10 Q8WOM1
 2.5 122 10 Q84PV3
 2.5 123 16 Q8EFG2
 2.5 123 16 Q8UGH2
 2.5 123 16 Q7U823
 2.5 124 10 Q7XRK7
 2.5 125 5 Q9UIP6

ALIGNMENTS

PRELIMINARY; PRT; 330 AA.
 TREMBLrel. 23, Created
 TREMBLrel. 23, Last sequence update
 TREMBLrel. 25, Last annotation update
 (Human).
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Heria; Primates; Catarrhini; Hominidae; Homo.
 106;

1 N.A.
 1924; PubMed=12411489;
 B.; Medema J.P.; Lopez-Fraga M.; Lozano J.C.;
 M.; Picard A.; Martinez A.C.; Garcia-Sanz J.A.,
 is hybrid mRNA encodes TWE-PRIL, a functional cell surface
 fusion protein."
 11-5720(2002).
 1; AAL90443.1; -.
 0; C.membrane; IEA.
 4; F.tumor necrosis factor receptor binding; IEA.
 5; P.immune response; IEA.
 106052; TNF family.
 1008983; TNF_like.
 1; TNF; 1.
 17; TNF; 1.
 1251; TNF 1; 1.
 1049; TNF 2; 2.
 10 AA; 36588 MW; FC6F3BCA29C029AE CRC64;

58.5%; Score 166; DB 4; Length 330;
 arity 100.0%; Pred. No. 1.3e-155;
 conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RSQRGRGEGPTALLVPLALGLALACGLLLAVSLGSRASLSQAQEEL 95
 RSQRGRGEGPTALLVPLALGLALACGLLLAVSLGSRASLSQAQEEL 60
 RSQRGRGEGPTALLVPLALGLALACGLLLAVSLGSRASLSQAQEEL 155
 RSQRGRGEGPTALLVPLALGLALACGLLLAVSLGSRASLSQAQEEL 120
 RSQRGRGEGPTALLVPLALGLALACGLLLAVSLGSRASLSQAQEEL 201
 RSQRGRGEGPTALLVPLALGLALACGLLLAVSLGSRASLSQAQEEL 166

PRELIMINARY; PRT; 410 AA.
 (TREMBLrel. 23, Created)
 (TREMBLrel. 23, Last sequence update)
 (TREMBLrel. 25, Last annotation update)

DE Tumor necrosis factor.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotati
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK044387; BAC31897.1; -.
 DR PIR; PT0714; PT0714.
 DR GO; GO:0016020; C.membrane; IEA.
 DR GO; GO:0005184; F.tumor necrosis factor receptor binding; IEA.
 DR GO; GO:0006955; P.immune response; IEA.
 DR InterPro; IPR006052; TNF family.
 DR InterPro; IPR008983; TNF_like.
 DR SMART; SM00207; TNF; 2.
 DR PROSITE; PS00251; TNF 1; 1.
 DR PROSITE; PS0049; TNF 2; 2.
 SQ SEQUENCE 410 AA; 45881 MW; 590A4B74C33FB8D4 CRC64;

Query Match 11.3%; Score 32; DB 11; Length 410;
 Best Local Similarity 100.0%; Pred. No. 8.9e-23;
 Matches 32; Conservative 0; Mismatches 0; Indels 0;

QY 139 RRAIAHYEHVPRPGQDGAQAGVDGTVSGWEE 170
 DB 105 RRAIAHYEHVPRPGQDGAQAGVDGTVSGWEE 136

RESULT 3

Q7VVB7
 ID Q7VVB7 PRELIMINARY; PRT; 438 AA.
 AC Q7VVB7;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Putative chloride-channel protein.
 GN BP2760.
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OC NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J.; Sebahia M.; Preston A.; Murphy L.D.; Thomson N.;
 RA Harris D.E.; Holden M.T.G.; Churcher C.M.; Bentley S.D.; Mungall
 RA Cerdano-Tarraga A.M.; Temple L.; James K.; Harris B.; Quail M.A.;
 RA Achtman M.; Ackin R.; Baker S.; Basham D.; Bason N.; Cherevach I.
 RA Chillingworth T.; Collins M.; Cronin A.; Davis P.; Doggett J.;
 RA Felwell T.; Goble A.; Hamlin N.; Hauser H.; Holroyd S.; Jagels I
 RA Leather S.; Moule S.; Norberczak H.; O'Neil S.; Ormond D.; Price
 RA Rabinowitch E.; Rutter S.; Sanders M.; Saunders D.; Seeger K.;
 RA Sharp S.; Simmonds M.; Skelton J.; Squares R.; Squares S.; Stevel
 RA Unwin L.; Whitehead S.; Barrell B.G.; Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pert
 RT Bordetella parapertussis and Bordetella bronchiseptica."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; EX640419; CAE43035.1; -.
 KW Complete proteome.

SQ SEQUENCE 438 AA; 45402 MW; AD51A65B59599D8 CRC64;

Query Match 4.2%; Score 12; DB 16; Length 438;
 Best Local Similarity 100.0%; Pred. No. 0.0062;
 Matches 12; Conservative 0; Mismatches 0; Indels 0;

ALACLGIL 73
|||||
ALACLGIL 285

RELIMINARY; PRT; 470 AA.

TrEMBLrel. 25, Created)
TrEMBLrel. 25, Last sequence update)
TrEMBLrel. 25, Last annotation update)
side-channel protein.

apertusis.
eobacteria; Betaproteobacteria; Burkholderiales;
; Bordetella.

N.A.

/ ATCC BAA-587;
954; PubMed=12910271;
enathia M., Preston A., Murphy L.D., Thomson N.,
olden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
ia A.M., Temple L., James K., Harris B., Quail M.A.,
kin R., Baker S., Basham D., Bason N., Cherevach I.,
T., Collins M., Cronin A., Davis P., Doggett J.,
oble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
ule S., Norberczak H., O'Neill S., Ormond D., Price C.,
E., Rutter S., Sanders M., Saunders D., Seeger K.,
onds M., Skelton J., Squares R., Squares S., Stevens K.,
head S., Barrell B.G., Maskell D.J.;
analysis of the genome sequences of Bordetella pertussis,
apertusis and Bordetella bronchiseptica.";
3:32-40(2003).
); CAB37862.1; -.
); AA; 49233 MW; 5B92DD05E920BDSF CRC64;

4.28; Score 12; DB 16; Length 470;

arity 100.0%; Pred. No. 0.066;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

ALACLGIL 73
|||||
ALACLGIL 317

RELIMINARY; PRT; 111 AA.

TrEMBLrel. 20, Created)
TrEMBLrel. 20, Last sequence update)
TrEMBLrel. 24, Last annotation update)
protein z3516.

oli O157:H7.
eobacteria; Gammaproteobacteria; Enterobacteriales;
aceae; Escherichia.
334;

N.A.

7 / EDL933 / ATCC 700927;
935; PubMed=11206551;
unkett G. III, Burland V., Mau B., Glaesner J.D.,
yeh G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
ckett J., Klink A., Boutin A., Shao Y., Miller L.,
Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
nantharaman T.S., Lin J., Yen G., Schwartz D.C.,
latner F.R.;
nce of enterohaemorrhagic Escherichia coli O157:H7.";
9-533(2001).

DR EMBL; AR005458; AAG57389.1; -.
DR PIR; A85866; A85866.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 111 AA; 12165 MW; 7CEFC93D786CD759 CRC64;

Query Match 3.5%; Score 10; DB 16; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; G

QY 63 LGLALACLGIL 72
|||||
Db 40 LGLALACLGIL 49

RESULT 6

Q8FFL8 PRELIMINARY; PRT; 111 AA.
AC Q8FFL8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN C2800.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Weich R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome seque
of uropathogenic Escherichia coli".
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016763; AA081254.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 111 AA; 12196 MW; C0A977B6F77A4B87 CRC64;

Query Match 3.5%; Score 10; DB 16; Length 111;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; G

QY 63 LGLALACLGIL 72
|||||
Db 40 LGLALACLGIL 49

RESULT 7

Q7UC61 PRELIMINARY; PRT; 111 AA.
AC Q7UC61;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sucrose-6 phosphate hydrolase.
GN S2567.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,

1 N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
Blattner F.R.;
ome sequence and comparative genomics of Shigella
type 2a strain 2457T.",
1. 71:2775-2786(2003).
36; AAP17670.1; -.

11 AA; 12224 MW; 7CFA06CC46A32672 CRC64;

3.5%; Score 10; DB 16; Length 111;
larity 100.0%; Pred. No. 0.17; 0; Indels 0; Gaps 0;
Conservative 0; Mismatches 0; Mismatches 0; Indels 0;

ALACGL 72

|||||

ALACGL 49

PRELIMINARY; PRT; 142 AA.

(TREMBLrel. 18, Created)

(TREMBLrel. 18, Last sequence update)

(TREMBLrel. 18, Last annotation update)

asport protein.

ti (Mesorhizobium loti).

teobacteria; Alphaproteobacteria; Rhizobiales;

iaceae; Mesorhizobium.

91;

M N.A.

03099;

2930; PubMed=11214968;

akamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,

Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,

Yanada M., Tabata S.;

name structure of the nitrogen-fixing symbiotic bacterium

m loti";

31-338(2000).

99; BAB49526.1; -.

teome.

42 AA; 14884 MW; ODCA7842C85A5B6F CRC64;

3.2%; Score 9; DB 16; Length 142;

larity 100.0%; Pred. No. 2.1; 0; Indels 0; Gaps 0;

Conservative 0; Mismatches 0; Mismatches 0; Indels 0;

LVPLAL 61

|||||

LVPLAL 130

PRELIMINARY; PRT; 749 AA.

(TREMBLrel. 25, Created)

(TREMBLrel. 25, Last sequence update)

(TREMBLrel. 25, Last annotation update)

I PeaB protein.

769.

cus marinus (strain MIT 9313).

anobacteria; Prochlorophytes; Prochlorococcaceae;

cus.

4547;

M N.A.

5698; PubMed=12917642;

zimer F.W., Lamerdin J., Malfatti S., Chain P.,

RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regalia W., Shah M.
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;

"Genome divergence in two Prochlorococcus ecotypes reflects ocea

RT niche differentiation";

RL Nature 424:1042-1047(2003).

DR EMBL; BX572100; CAE21944.1; -.

KW Photosystem I; Complete proteome.

SQ SEQUENCE 749 AA; 83231 MW; B1D496645F1C790C CRC64;

Query Match 3.2%; Score 9; DB 16; Length 749;

Best Local Similarity 100.0%; Pred. No. 9.4; 0; Indels 0;

Matches 9; Conservative 0; Mismatches 0; Mismatches 0;

QY 63 LGLALACLG 71

|||||

Db 347 LGLALACLG 355

RESULT 10

Q8PPZ8

ID Q8PPZ8 PRELIMINARY; PRT; 766 AA.

AC Q8PPZ8;

DT 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE C-type cytochrome biogenesis protein (Copper tolerance).

GN DSBD OR XAC0534.

OS Xanthomonas axonopodis (pv. citri).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xanthomonas.

OX NCBI_TaxID=92829;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=306 / ATCC 13902 / XV 101;

RX MEDLINE=22022145; PubMed=12024217;

da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.

Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N

Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,

RA Camarotte G., Cannavari F., Cardozo J., Chamberg F., Clapina L.P

Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H

Formiglieri E.F., Franco M.C., Greggio C.C., Gruber A.,

RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.

Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.

RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,

Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V

Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,

RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R

Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,

RA Setubal J.C., Kitajima J.P.;

RT "Comparison of the genomes of two Xanthomonas pathogens with dif

host specificities";

RL Nature 417:459-463(2002).

DR EMBL; AF011680; AM35423.1; -.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005489; F:electron transporter activity; IEA.

DR GO; GO:0017004; P:cytochrome biogenesis; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR003834; Cytococh TM.

DR InterPro; IPR006662; ThioRed.

DR InterPro; IPR006663; ThioRedox_dom2.

DR Pfam; PF02683; DSBD; 1.

DR PROSITE; PS00194; THIOREDOXIN; 1.

KW Complete proteome.

SQ SEQUENCE 766 AA; 81014 MW; 3A1955A07DB8A9CA CRC64;

Query Match 3.2%; Score 9; DB 16; Length 766;

Best Local Similarity 100.0%; Pred. No. 9.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0;

QY 59 LAIGLGLAL 67

|||||
GLAL 542

RELIMINARY; PRT; 1208 AA.

TrEMBLrel. 25, Created)
TrEMBLrel. 25, Last sequence update)
TrEMBLrel. 25, Last annotation update)
; multi-functional protein.

ibaltica.
ctomycetes; Planctomycetacia; Planctomycetales;
eae; Pirellula.

N.A.

113; PubMed=12835416;
; Kube M., Bauer M., Teeling H., Lombardot T.,
ie D., Beck A., Borzym K., Heitmann K., Rabus R.,
Anann R., Reinhardt R.;
ome sequence of the marine planctomycete Pirellula sp.

ad. Sci. U.S.A. 100:8298-8303 (2003).

; CAD76265.1; -.

18 AA; 132047 MW; OFPE225741021E8C CRC64;

3.2%; Score 9; DB 16; Length 1208;

arity 100.0%; Pred. No. 15;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

IQEPA 91

IQEPA 42

RELIMINARY; PRT; 143 AA.

TrEMBLrel. 16, Created)
TrEMBLrel. 16, Last sequence update)
TrEMBLrel. 24, Last annotation update)

SP. (strain NRC-1 / ATCC 700922 / JCM 11081).

archaeota; Halobacteria; Halobacteriales;

ae; Halobacterium.

391;

N.A.

483; PubMed=11016950;
edy S.P., Mahairas G.G., Berquist B., Pan M.,
asky S.R., Balliga N.S., Thorsson V., Shrogha J.,
Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
Keller K., Cruz R., Danson M.J., Hough D.W.,
Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
A., Peck R.F., Pohlchröder M., Spudich J.L., Jung K.-H.,
tas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
owe T.M., Liang P., Riley M., Hood L., Dassarma S.;
nce of Halobacterium species NRC-1.";
cad. Sci. U.S.A. 97:12176-12181 (2000).

6; AAG18715.1; -.

G84168.

006976; Vanz.

; Vanz; 1.

3 AA; 15648 MW; 45466B6328EF3468 CRC64;

Query Match 2.8%; Score 8; DB 17; Length 143;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 61 LGLGLALA 68
Db 55 LGLGLALA 62

RESULT 13

Q9SD11

ID Q9SD11 PRELIMINARY; PRT; 151 AA.

AC Q9SD11;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Hypothetical protein (OSUNBA0036E02.6 protein) (B1085F09.2

DE protein).

GN B1085F09.2.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC

clone:P0003H10.";

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC

clone:OSUNBA0036E02.";

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC

clone:B1085F09.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP000815; BAA87834.1; -.

DR EMBL; AP002862; BAB17732.1; -.

DR EMBL; AP003103; BAB44106.1; -.

DR Gramene; Q9SD11; -.

SQ SEQUENCE 151 AA; 16632 MW; EC68451ECA2BD71D CRC64;

Query Match 2.8%; Score 8; DB 10; Length 151;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 43 RRRGRGGE 50
Db 131 RRRGRGGE 138

RESULT 14

Q848K4

ID Q848K4 PRELIMINARY; PRT; 154 AA.

AC Q848K4;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein (Fragment).

OS Gamma-proteobacterium Hot 75m4.

OG Plasmid pAK106.

OC Bacteria; environmental samples.

OX NCBI_TaxID=77133;

RN [1]


```

1 N.A.
2 661; PubMed=12620823;
3 Waschkowitz T., Bowien S., Henne A., Daniel R.;
4 and Screening of Metagenomic Libraries Derived from
5 itures: Generation of a Gene Bank for Genes Conferring
6 reductase Activity on Escherichia coli.;
7 1. Microbiol. 69:1408-1416(2003).
8 30; AAO59972.1; -.
9 21; C:extrachromosomal DNA; IEA.
10 protein; Plasmid.
11 1
12 34 AA; 16234 MW; 3A89A072D5B7E137 CRC64;
13 2.8%; Score 8; DB 2; Length 154;
14 larity 100.0%; Pred. No. 22;
15 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
16 {GRRG 49
17 |||||
18 {GRRG 119
19
20 PRELIMINARY; PRT; 190 AA.
21 (TrEMBLrel. 23, Created)
22 (TrEMBLrel. 23, Last sequence update)
23 (TrEMBLrel. 24, Last annotation update)
24 protein.
25 s agalactiae (serotype III).
26 rmicutes; Lactobacillales; Streptococcaceae;
27 s.
28 16495;
29
30 N.A.
31 5 / Serotype III;
32 2508; PubMed=12354221;
33 usniok C., Buchrieser C., Chevallier F., Frangeul L.,
34 guine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
35 ence of Streptococcus agalactiae, a pathogen causing
36 natal disease.;"
37 31. 45:1499-1513(2002).
38 49; CAB46822.1; -.
39 s1163; -.
40 8008172; Adenylate_cyc.
41 8; CYTH; 1.
42 protein; Complete proteome.
43 90 AA; 22178 MW; AB2AD33C2CB6FBF3 CRC64;
44 2.8%; Score 8; DB 16; Length 190;
45 larity 100.0%; Pred. No. 27;
46 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
47 IRTLP 265
48 |||||
49 IRTLP 57
50
51 PRELIMINARY; PRT; 190 AA.
52 (TrEMBLrel. 23, Created)
53 (TrEMBLrel. 23, Last sequence update)
54 (TrEMBLrel. 24, Last annotation update)
55 pothetical protein.
56 s agalactiae (serotype V).
57 rmicutes; Lactobacillales; Streptococcaceae;
58 s.
59
60 NCBI_TaxID=216466;
61 [1]
62 SEQUENCE FROM N.A.
63 RC STRAIN=2603 V/R / Serotype V;
64 MEDLINE=22222988; PubMed=12200547;
65 Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson
66 Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
67 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C
68 DeBoy R.T., Durkin A.S., Klonay J.F., Madupu R., Lewis M.R.,
69 Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
70 Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mo;
71 Iacobini E.T., Bretttoni C., Galli G., Mariani M., Vegni F., Maior
72 Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
73 Fraser C.M.;
74 "Complete genome sequence and comparative genomic analysis of an
75 emerging human pathogen, serotype V Streptococcus agalactiae.;"
76 Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
77 EMBL; AE014242; AAM99977.1; -.
78 TIGR; SAG1096; -.
79 InterPro; IPR008172; Adenylate_cyc.
80 Pfam; PF01928; CYTH; 1.
81 KW Hypothetical protein; Complete proteome.
82 SQ SEQUENCE 190 AA; 22178 MW; AB2AD33C2CB6FBF3 CRC64;
83
84 Query Match 2.8%; Score 8; DB 16; Length 190;
85 Best Local Similarity 100.0%; Pred. No. 27;
86 Matches 8; Conservative 0; Mismatches 0; Indels 0;
87
88 QY 258 SLRIRTLTP 265
89 |||||
90 Db 50 SLRIRTLTP 57
91
92 RESULT 17
93 Q8L4K2 PRELIMINARY; PRT; 193 AA.
94 AC Q8L4K2;
95 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
96 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
97 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
98 DE Hypothetical protein.
99 GN OSJNBA0079H13.8 OR OSJNBA0038H12.21.
100 OS Oryza sativa (japonica cultivar-group).
101 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyt
102 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
103 OC Ehrhartoideae; Oryzaceae; Oryza.
104 ON NCBI_TaxID=39947;
105 RN [1]
106 RP SEQUENCE FROM N.A.
107 RC STRAIN=cv. Nipponbare;
108 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
109 Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
110 Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
111 Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
112 Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C
113 "Oryza sativa chromosome 10 BAC OSJNBA0079H13 genomic sequence."
114 Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
115 [2]
116 RN SEQUENCE FROM N.A.
117 RP STRAIN=cv. Nipponbare;
118 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M
119 Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.
120 Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S
121 Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
122 Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
123 White O., Salzberg S.L., Fraser C.M.;
124 "Oryza sativa chromosome 10 BAC OSJNBA0038H12 genomic sequence."
125 Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
126 [3]
127 RN SEQUENCE FROM N.A.
128 RP STRAIN=cv. Nipponbare;
129 RC The Rice Chromosome 10 Sequencing Consortium;
130 RA "In-depth view of structure, activity, and evolution of rice

```

```

";
66-1569 (2003).
N.A.
ponbare;
ng R.A., McCombie W.R., Messing J., Yuan Q.;
-2003) to the EMBL/GenBank/DBJ databases.
; AAM54153.1; -
; AAN04965.1; -
; AAP52530.1; -
2; -.
rotein.
AA; 20812 MW; 719544BFC9A0790 CRC64;
2.8%; Score 8; DB 10; Length 193;
urity 100.0%; Pred. No. 27; Indels 0; Gaps 0;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
IRGE 50
|||
RGE 123
ELIMINARY; PRT; 197 AA.
'EMBLrel. 12, Created)
'EMBLrel. 12, Last sequence update)
'EMBLrel. 24, Last annotation update)
rotein TM0469.
itima.
otogae; Thermotogales; Thermotogaceae; Thermotoga.
16;
N.A.
DSM 3109;
116; PubMed=10360571;
layton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Key E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
Jeterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Smith H.O., Venter J.C., Fraser C.M.;
lateral gene transfer between Archaea and Bacteria from
e of Thermotoga maritima.";
1-329(1999).
; AAD35553.1; -.
; 272374.
rotein; Complete proteome.
/ AA; 22919 MW; 41E2C8E3C09180EC CRC64;
2.8%; Score 8; DB 16; Length 197;
urity 100.0%; Pred. No. 28; Indels 0; Gaps 0;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
EIS 9
|||
EIS 142
REIMINARY; PRT; 211 AA.
'EMBLrel. 13, Created)
'EMBLrel. 22, Last sequence update)
'EMBLrel. 24, Last annotation update)
in.
lanogaster (Fruit fly).

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin I
RA Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dum
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman
RA Fostler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum I
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng ;
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith ;
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2195-2195 (2000).
RN [2]
SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,
RA Gonzalez C., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam M., Jallali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Prounenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell ;
RA Hradercky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emert D., Frise E., de Grey A., Harris N.
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";

```

R-2000) to the EMBL/GenBank/DBJ databases.

[N.A.
Slinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
R-2000) to the EMBL/GenBank/DBJ databases.

[N.A.

P-2002) to the EMBL/GenBank/DBJ databases.

7; RAF49452.2; -

0036638; CGI3033.

004011; Gyr.

004019; YLP_motif.

; Gyr; 1.

; YLP; 5.

1 AA; 23779 MW; D2554983E91F5107 CRC64;

2.8%; Score 8; DB 5; Length 211;
arity 100.0%; Pred. No. 29;
conservative 0; Mismatches 0; Indels 0; Gaps 0;

WVSL 79

||||

WVSL 16

PRELIMINARY; PRT; 211 AA.

TREMBLrel. 21, Created)

TREMBLrel. 21, Last sequence update)

TREMBLrel. 24, Last annotation update)

Planogaster (Fruit fly).

stazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

opterygota; Diptera; Brachycera; Muscomorpha;

Drosophilidae; Drosophila.

27;

[N.A.

ey;

Brokstein P., Hong L., Agbayani A., Carlson J.,

avez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

nzalez M., Guarin H., Kronmiller B., Li P., Liao G.,

ungall C.J., Nunoo J., Pachter J., Faragas V., Park S.,

mananavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

AR-2002) to the EMBL/GenBank/DBJ databases.

4; AAL90372.1; -

0063673; BCDNA.RE50345.

004011; Gyr.

004019; YLP_motif.

; Gyr; 1.

; YLP; 5.

11 AA; 23780 MW; 82FF4983E91F510A CRC64;

2.8%; Score 8; DB 5; Length 211;
arity 100.0%; Pred. No. 29;
conservative 0; Mismatches 0; Indels 0; Gaps 0;

WVSL 79

||||

WVSL 16

PRELIMINARY; PRT; 220 AA.

(TREMBLrel. 23, Created)

(TREMBLrel. 23, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Membrane protein, putative.

GN BRA0991.

OS Brucella suis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Brucellaceae; Brucella.

OX NCBI_TaxID=29461;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=1330 / Biovar 1;

RX MEDLINE=22247741; PubMed=12271122;

RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.

RA Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,

RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,

RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken

RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,

RA Hoover D.L., Lindler L.E., Hailing S.M., Boyle S.M., Fraser C.M.,

RT "The Brucella suis genome reveals fundamental similarities between

RT animal and plant pathogens and symbionts.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).

DR EMBL; AE014592; AAN34160.1; -

DR TIGR; BRA0991; -

DR InterPro; IPR007916; UPF0191.

DR Pfam; PF05252; UPF0191; 1.

KW Complete proteome.

SQ SEQUENCE 220 AA; 24796 MW; AC2C060433169497 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 220;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QV 54 ALLVPLAL 61

|||||

Db 132 ALLVPLAL 139

RESULT 22

Q89GW9

AC Q89GW9; PRELIMINARY; PRT; 232 AA.

DT 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE BLR6226 protein.

GN BLR6226.

OS Bradyrhizobium japonicum.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Bradyrhizobium.

OX NCBI_TaxID=375;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=USDA 110;

RX MEDLINE=22484998; PubMed=12597275;

RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,

RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,

RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada

RA Tabata S.;

RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium

RT Bradyrhizobium japonicum USDA110.";

RL DNA Res. 9:189-197 (2002).

DR EMBL; AF005957; BAC51491.1; -

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000515; BPD_transp.

DR Pfam; PF00528; BPD_transp; 1.

KW Complete proteome.

SQ SEQUENCE 232 AA; 23704 MW; CD805BD1F43F1B46 CRC64;

Query Match

Best Local Similarity 100.0%; Pred. No. 32;

Matches 8; Conservative 0; Mismatches 0; Indels 0;

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou
 RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sajer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL; AF082575; AAC98784.1; --
 DR EMBL; AF084866; AAG07909.1; --
 DR FIR; D83080; D83080.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 KW Transmembrane; Complete proteome.
 SQ SEQUENCE 278 AA; 30793 MW; C623F1AB0691CPEF CRC64;

 Query Match 2.8%; Score 8; DB 16; Length 278;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G

 QY 70 LGLLLAVV 77
 Db 47 LGLLLAVV 54

 RESULT 25
 Q82J43 PRELIMINARY; PRT; 278 AA.
 ID Q82J43
 AC Q82J43;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative metalloproteinase.
 GN SAV2939.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industr.
 RT microorganism Streptomyces avermitilis."
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AF085033; BAC70650.1; --
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0008237; F:metallopeptidase activity; IEA.
 DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000013; Peptidase_M7.
 DR Pfam; PF02031; Peptidase_M7; 1.
 DR PRINTS; PR00787; NEUTRALPTASE.
 DR ProDom; PD016028; Peptidase_M7; 1.
 KW Complete proteome.
 SQ SEQUENCE 278 AA; 28113 MW; 9545813BCAC0BFA2 CRC64;

 Query Match 2.8%; Score 8; DB 16; Length 278;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G

 QY 59 LALGLGLIA 66

 LGL 65
 LGL 209

 LELIMINARY; PRT; 276 AA.
 TREMBLrel. 24, Created)
 TREMBLrel. 24, Last sequence update)
 TREMBLrel. 25, Last annotation update)

 aliana (Mouse-ear cress).
 idiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Magnoliophyta; eudicotyledons; core eudicots; rosids;
 Brassicales; Brassicaceae; Arabidopsids.
 2;

 N.A.
 H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 ashizaki Y., Huan V.W., Ishida J., Jones T., Kamiya A.,
 G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
 dera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 wick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
 zaki K., Davis R.W., Theologis A., Ecker J.R.;
 JRF clones."
 1-2003) to the EMBL/GenBank/DBJ databases.
 AAP21226.1; --
 F:electron transporter activity; IEA.
 P:electron transport; IEA.
 00345; Cytochrome BS.
 101841; ZnF ring.
 zf-C3HC4; 1.
 RING; 1.
 90; CYTOCHROME C; 1.
 189; ZF_RING; 2; 1.
 AA; 31368 MW; 99DE3DA0CB2C0BDF CRC64;

 2.8%; Score 8; DB 10; Length 276;
 100.0%; Pred. No. 37;
 0; Mismatches 0; Indels 0; Gaps 0;
 0; Mismatches 0; Indels 0; Gaps 0;

 HSA 10
 HSA 221

 LELIMINARY; PRT; 278 AA.
 TREMBLrel. 10, Created)
 TREMBLrel. 10, Last sequence update)
 TREMBLrel. 24, Last annotation update)
 Protein AMPE.

 ruginosa.
 eobacteria; Gammaproteobacteria; Pseudomonadales;
 ae; Pseudomonas.

 N.A.
 5692 / PA01;
 Dargis M., Huletsky A.;
 in *Pseudomonas aeruginosa* encodes a negative regulator
 lactanase expression."
 1-1998) to the EMBL/GenBank/DBJ databases.

 N.A.
 5692 / PA01;
 137; PubMed=10984043;
 pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,

M N.A.
/6J: TISSUE=Embrvo. and Embrvonic stem cells;

OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

rosophilidae; Drosophila.

7;

N.A.

06; PubMed-10731132;

iniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Lewis S.E., Li P.W., Hoskins R.A., Galle R.F.,
Rothberg L., Richards S., Ashburner M., Henderson S.N.,
Lortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Rogers Y.H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
E.C., Baxter E.G., Heit G., Nelson C.R., Miklos G.D.,
Bayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Lasus A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Espinosa P.V., Bertram B.P., Bhandari D., Bolshakov S.,
Chen M.R., Bouck J., Brokstein P., Brotter P.,
Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cawley S., Dahlke C., Davenport L.B., Davies P.,
DeLcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
P.L.E., Downes M., Dugan-Rocha S., Durkin B.C., Dunn P.,
Vangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Giannoulatou A.E., Garg N.S., Gelbart W.M., Glasser K.,
G.F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Joston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
B., McIntosh T.C., McLeod M.P., McPherson D.,
Kilshina N.V., Mobarry C., Morris J., Moshrefi A.,
My M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Wingman K., Saunders R.D.C., Scheeler F., Shen H.,
Ten-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Stilling A.C., Stapleton M., Strong R., Sun E.,
Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wasserman D.A., Weinstock G.M., Weissbach J.,
Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Long F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Venter E.W., Rubin G.M., Venter J.C.;
Sequence of Drosophila melanogaster.";
85-2195(2000).

N.A.

Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
H., Baldwin D., Banazon J., Beeson K.Y., Busam D.A.,
Center A., Champe M., Davenport L.B., Dietz S.M.,
Fetisov V., Galle R.F., Garg N.S., George R.A.,
Houck J., Hoskins R.A., Hostin D., Howland T.J.,
Irali M., Kruse D., Li P., Mattel B., Moshrefi A.,
Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
Pacis V., Park S., Patel S., Pfeiffer B.,
S., Pittman G.S., Puri V., Richards S., Scheeler F.,
Strong R., Svitskas R., Tector C., Tyler D.,
Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
Drosophila melanogaster genome.";
(2000) to the EMBL/GenBank/DBJ databases.

N.A.

by M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
Guang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
Tuan C., Berman B., Carlson J.W., Celniker S.E.,
Sdale R., Emmert D., Frise E., de Grey A., Harris N.,
Marshall B., Millburn G., Richter J., Russo S.,
Smith E., Shu S., Smutniak F., Whitfield E.,
Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
Drosophila melanogaster genome.";
(2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RA FlyBase;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN-Berkeley;

RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Pacle J., Park S.,

RA Patel S., Phuanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Celniker S.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AS003600; AAF51889.2; -;

DR EMBL; AY070982; AAL48604.1; -;

DR FlyBase; FBgn0037428; CGI119.

SQ SEQUENCE 306 AA; 34083 MW; 32B69371475A48F9 CRC64;

Query Match 2.8%; Score 8; DB 5; Length 306;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; G;

QY 56 LVPLALGL 63

Db 170 LVPLALGL 177

RESULT 29

Q864L1

ID Q864L1 PRELIMINARY; PRT; 317 AA.

AC Q864L1;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Melanocortin-1 receptor.

OS MC1R.

GN Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

OX NCBI_TaxID=9598;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=3;

RX MEDLINE-22572539; PubMed-12687585;

RA Mundy N.I., Kelly J.;

RT "Evolution of a pigmentation gene, the melanocortin-1 receptor, i

RT Primates.";

RL Am. J. Phys. Anthropol. 121:67-80(2003).

DR EMBL; AY205086; AAP30960.1; -;

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.

DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.

DR InterPro; IPR000276; GPCR_Rhodopsn.

DR Pfam; PF00001; 7tm.1.1.

DR PRINTS; PR00237; GPCRHOOPS.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.

SQ SEQUENCE 317 AA; 34710 MW; 8815D21464BD2475 CRC64;

Query Match 2.8%; Score 8; DB 6; Length 317;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; G;

QY 137 RARRATAA 144

Db 160 RARRATAA 167

PRELIMINARY; PRT; 317 AA.
 TREMBLrel. 24, Created)
 TREMBLrel. 24, Last sequence update)
 TREMBLrel. 25, Last annotation update)
 1 receptor.
 is (Orangutan).
 :tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 heria; Primates; Catarrhini; Hominidae; Pongo.
 ;00;
 1 N.A.
 2539; PubMed=12687585;
 Kelly J.;
 : a pigmentation gene, the melanocortin-1 receptor, in
 Anthropol. 121:67-80(2003).
 37; AAP30961.1; -.
 11; C:integral to membrane; IEA.
 72; F:receptor activity; IEA.
 14; F:rhodopsin-like receptor activity; IEA.
 16; P:G-protein coupled receptor protein signalin. . .; IEA.
 1000276; GPCR_Rhodpsn.
 ; 7tm 1; 1.
 737; GPCR_RHODPSN.
 1237; G_PROTEIN_RECEP_F1_1; 1.
 1262; G_PROTEIN_RECEP_F1_2; 1.
 17 AA; 34749 MW; 83D385B1B010D865 CRC64;
 2.8%; Score 8; DB 6; Length 317;
 larity 100.0%; Pred.No.42;
 Conservative 0; Mismatches 0; Indels 0; Gaps
 AIAA 144
 ||||
 AIAA 167
 PRELIMINARY; PRT; 317 AA.
 TREMBLrel. 24, Created)
 TREMBLrel. 24, Last sequence update)
 TREMBLrel. 25, Last annotation update)
 1 receptor.
 :rina (Pig-tailed macaque).
 :tazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 heria; Primates; Catarrhini; Cercopithecidae;
 :ae; Macaca.
 345;
 1 N.A.
 2539; PubMed=12687585;
 Kelly J.;
 f a pigmentation gene, the melanocortin-1 receptor, in
 Anthropol. 121:67-80(2003).
 99; AAP30973.1; -.
 21; C:integral to membrane; IEA.
 72; F:receptor activity; IEA.
 34; F:rhodopsin-like receptor activity; IEA.
 86; P:G-protein coupled receptor protein signalin. . .; IEA.
 R000276; GPCR_Rhodpsn.
 1; 7tm 1; 1.

```

N.A.
39; PubMed=12687585;
illy J.;
a pigmentation gene, the melanocortin-1 receptor, in
anthropol. 121:67-80(2003).
; AAP30976.1; -.
; C:integral to membrane; IEA.
; F:receptor activity; IEA.
; F:rhodopsin-like receptor activity; IEA.
; P:G-protein coupled receptor protein signalin. . . ; IEA.
00276; GPCR_Rhodpsn.
7tm.1; 1.
7; GPCR_Rhodopsn.
37; G_PROTEIN_RECEP_F1_1; 1.
62; G_PROTEIN_RECEP_F1_2; 1.
AA; 34779 MW; 1A091A65BDB8CBAC CRC64;
2.8%; Score 8; DB 6; Length 317;
rity 100.0%; Pred.No.42;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
IAAA 144
|||
IAAA 167

ELIMINARY; PRT; 317 AA.
TREMBlrel. 24, Created)
TREMBlrel. 24, Last sequence update)
TREMBlrel. 25, Last annotation update)
receptor.
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
.953;
N.A.
39; PubMed=12687585;
illy J.;
a pigmentation gene, the melanocortin-1 receptor, in
anthropol. 121:67-80(2003).
; AAP31003.1; -.
; C:integral to membrane; IEA.
; F:receptor activity; IEA.
; F:rhodopsin-like receptor activity; IEA.
; P:G-protein coupled receptor protein signalin. . . ; IEA.
00276; GPCR_Rhodpsn.
7tm.1; 1.
7; GPCR_Rhodopsn.
37; G_PROTEIN_RECEP_F1_1; 1.
62; G_PROTEIN_RECEP_F1_2; 1.
AA; 34654 MW; DA5F4420DFECC4B CRC64;
2.8%; Score 8; DB 6; Length 317;
rity 100.0%; Pred.No.42;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
IAAA 144
|||
IAAA 167

```

```

RESULT 35
Q864G7
ID Q864G7 PRELIMINARY; PRT; 317 AA.
AC Q864G7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Melanocortin-1 receptor.
GN MCLR.
OS Ateles paniscus (Black spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ate.
OX NCBI_TaxID=9510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX MEDLINE=22572539; PubMed=12687585;
RA Mundy N.I., Kelly J.;
RT "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
RT primates.";
RL Am. J. Phys. Anthropol. 121:67-80(2003).
DR EMBL; AY205130; AAP31004.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 317 AA; 34719 MW; 5481D6A1B9085D43 CRC64;

Query Match 2.8%; Score 8; DB 6; Length 317;
Best Local Similarity 100.0%; Pred.No.42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 RARRAIAA 144
|||
DB 160 RARRAIAA 167

RESULT 36
Q864G6
ID Q864G6 PRELIMINARY; PRT; 317 AA.
AC Q864G6;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Melanocortin-1 receptor.
GN MCLR.
OS Alouatta seniculus (Red howler monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
OC Alouatta.
OX NCBI_TaxID=9503;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=471;
RX MEDLINE=22572539; PubMed=12687585;
RA Mundy N.I., Kelly J.;
RT "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
RT primates.";
RL Am. J. Phys. Anthropol. 121:67-80(2003).
DR EMBL; AY205131; AAP31005.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.

```



```

DF 237; G_PROTEIN_RECEP_F1_1; 1.
DE 1262; G_PROTEIN_RECEP_F1_2; 1.
KW 7 AA; 34830 MW; 87F7EFAE347671E4 CRC64;
SQ 2.8%; Score 8; DB 6; Length 317;
    arity 100.0%; Pred. No. 42;
    conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY RAIAA 144
DE |||||
SQ RAIAA 167

RE PRELIMINARY; PRT; 317 AA.
DE (TrEMBLrel. 24, Created)
DE (TrEMBLrel. 24, Last sequence update)
DE (TrEMBLrel. 25, Last annotation update)
DE -1 receptor.
DE 1 (Bolivian red howler monkey).
DE stazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Theria; Primates; Platyrrhini; Cebidae; Alouattinae;
DE ||1123;
DE 1 N.A.
DE 1539; PubMed=12687585;
DE Kelly J.;
DE a pigmentation gene, the melanocortin-1 receptor, in
DE Anthropol. 121:67-80(2003).
DE 13; AAP31007.1; -.
DE 21; C:integral to membrane; IEA.
DE 72; F:receptor activity; IEA.
DE 34; F:rhodopsin-like receptor activity; IEA.
DE 36; P:G-protein coupled receptor protein signalin...; IEA.
DE {000276; GPCR_Rhodopsin.
DE 1; 7tm 1; 1.
DE 237; GPCR_Rhodopsin.
DE 237; G_PROTEIN_RECEP_F1_1; 1.
DE 1262; G_PROTEIN_RECEP_F1_2; 1.
DE 17 AA; 34686 MW; BA7E14APEC7EA971 CRC64;
DE 2.8%; Score 8; DB 6; Length 317;
DE arity 100.0%; Pred. No. 42;
DE conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY RAIAA 144
DE |||||
SQ RAIAA 167

RE PRELIMINARY; PRT; 317 AA.
DE (TrEMBLrel. 24, Created)
DE (TrEMBLrel. 24, Last sequence update)
DE (TrEMBLrel. 25, Last annotation update)
DE -1 receptor.
DE aya (Black howler monkey).
DE stazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Theria; Primates; Platyrrhini; Cebidae; Alouattinae;
DE ||1123;
DE 502;

RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=22572539; PubMed=12687585;
RA Mundy N.I., Kelly J.;
RT "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
    Primates.";
RL Am. J. Phys. Anthropol. 121:67-80(2003).
DR EMBL; AY205134; AAP31009.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 317 AA; 34692 MW; 91320E374CDB75DB CRC64;
    Query Match 2.8%; Score 8; DB 6; Length 317;
    Best Local Similarity 100.0%; Pred. No. 42;
    Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 137 RARRAIAA 144
DE |||||
DE 160 RARRAIAA 167

RESULT 39
Q864G2 PRELIMINARY; PRT; 317 AA.
AC Q864G2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Melanocortin-1 receptor.
GN MC1R.
OS Alouatta palliata (Mantled howler monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
OC Alouatta.
OX NCBI_TaxID=30589;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22572539; PubMed=12687585;
RA Mundy N.I., Kelly J.;
RT "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
    Primates.";
RL Am. J. Phys. Anthropol. 121:67-80(2003).
DR EMBL; AY205135; AAP31009.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 317 AA; 34670 MW; D5414E350E2435DE CRC64;
    Query Match 2.8%; Score 8; DB 6; Length 317;
    Best Local Similarity 100.0%; Pred. No. 42;
    Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 137 RARRAIAA 144
DE |||||
DE 160 RARRAIAA 167

```

```

>ELIMINARY; PRT; 317 AA.
>EMBLrel. 24, Created)
>EMBLrel. 24, Last sequence update)
>EMBLrel. 25, Last annotation update)
> receptor.

L.
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Platyrrhini; Cebidae; Alouattinae;
1253;
N.A.
339; PubMed=12687585;
illy J.;
a pigmentation gene, the melanocortin-1 receptor, in
anthropol. 121:67-80(2003).
; AAP31010.1; -.
; C: integral to membrane; IEA.
; F: receptor activity; IEA.
; F: rhodopsin-like receptor activity; IEA.
; P: G-protein coupled receptor protein signalin. . . ; IEA.
000276; GPCR_Rhodopsn.
; 7tm1; 1.
37; GPCR_Rhodopsn.
137; G_PROTEIN_RECEP_F1_1; 1.
162; G_PROTEIN_RECEP_F1_2; 1.
; AA; 34728 MW; 976BB14FF98B4966 CRC64;
2.8%; Score 8; DB 6; Length 317;
rity 100.0%; Pred. No. 42;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
IAA 144
IAA 167
>ELIMINARY; PRT; 317 AA.
>EMBLrel. 24, Created)
>EMBLrel. 24, Last sequence update)
>EMBLrel. 25, Last annotation update)
> receptor.

ata rubra.
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Strepsirhini; Lemuridae; Varecia.
04;
N.A.
339; PubMed=12687585;
illy J.;
a pigmentation gene, the melanocortin-1 receptor, in
anthropol. 121:67-80(2003).
; AAP31013.1; -.
; C: integral to membrane; IEA.
; F: receptor activity; IEA.
; F: rhodopsin-like receptor activity; IEA.
; P: G-protein coupled receptor protein signalin. . . ; IEA.
000276; GPCR_Rhodopsn.
; 7tm1; 1.
37; GPCR_Rhodopsn.

```

```

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
KW PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 317 AA; 34714 MW; C1F5DA35032717D7 CRC64;

Query Match 2.8%; Score 8; DB 6; Length 317;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 137 RARRATAA 144
Db 160 RARRATAA 167

RESULT 42
Q864F7 PRELIMINARY; PRT; 317 AA.
AC Q864F7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Melanocortin-1 receptor.
GN MCLR.
OS Varecia variegata variegata.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Varecia.
OX NCBI_TaxID=87289;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22572539; PubMed=12687585;
RA Mundy N.I., Kelly J.;
RT "Evolution of a pigmentation gene, the melanocortin-1 receptor, in
RT primates.";
RL Am. J. Phys. Anthropol. 121:67-80(2003).
DR EMBL; AX205140; AAP31014.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0004872; F: receptor activity; IEA.
DR GO; GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. .
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 317 AA; 34714 MW; C1F5DA35032717D7 CRC64;

Query Match 2.8%; Score 8; DB 6; Length 317;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 137 RARRATAA 144
Db 160 RARRATAA 167

RESULT 43
Q864F6 PRELIMINARY; PRT; 317 AA.
AC Q864F6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Melanocortin-1 receptor.
GN MCLR.
OS Eulemur fulvus (brown lemur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Eulemur.
OX NCBI_TaxID=13515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;

```

```

AC Q864F4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Melanocortin-1 receptor.
DE MCLR.
GN Lemur catta (Ring-tailed lemur).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Primates; Strepsirhini; Lemnidae; Lemur.
OX NCBI_TaxID=9447;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX MEDLINE=22572539; PubMed=12687585;
RA Mundy N.I., Kelly J.;
RT "Evolution of a pigmentation gene, the melanocortin-1 receptor,
RT primates.";
RL Am. J. Phys. Anthropol. 121:67-80(2003).
DR EMBL; AY205143; AAP31017.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; C:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 317 AA; 34667 MW; 3E7419FDEC2DE738 CRC64;

Query Match 2.8%; Score 8; DB 6; Length 317;
Best Local Similarity 100.0%; Pred.No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 137 RARRAIAA 144
DB 160 RARRAIAA 167

RESULT 46
OS8554 PRELIMINARY; PRT; 339 AA.
ID OS8554
AC OS8554;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PH0824.
GN PH0824.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococca
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Naga
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku
RA Fundahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguci
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hype:
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AF000003; BAA29917.1; -.
DR FIC; C71132; C71132.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008508; F:bile acid:sodium symporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR002657; BilAC/Na_symport.
DR Pfam; PF01758; SBF; 1.

```

rotein; Complete proteome.
AA; 37228 MW; E91697D5C8C3705F CRC64;

2.8%; Score 8; DB 17; Length 339;
rity 100.0%; Pred. No. 45;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

AWV 77

|||
AVV 118

ELIMINARY; PRT; 342 AA.

REMBLrel. 03, Created)
REMBLrel. 03, Last sequence update)
REMBLrel. 24, Last annotation update)

Human).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
6;

N.A.

28; PubMed-8702217;
on N.W., Liebert M., Grossman H.B.;
lysis of a gene, Bb1, overexpressed in bladder and
ma.";

AAB37433.1; -.

04299; MBOAT_fam.

MBOAT; 1.

AA; 38163 MW; 2B479EA8CFF1B91C CRC64;

2.8%; Score 8; DB 4; Length 342;
rity 100.0%; Pred. No. 45;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

HAL 67

|||
HAL 317

ELIMINARY; PRT; 343 AA.

REMBLrel. 17, Created)
REMBLrel. 17, Last sequence update)
REMBLrel. 22, Last annotation update)
rotein.

Human).

azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
6;

N.A.

and Colon;

1-2001) to the EMBL/GenBank/DBJ databases.

AAH03164.1; -.

AAH02512.1; -.

04299; MBOAT_fam.

MBOAT; 1.

rotein.

AA; 38727 MW; F71E7DBF74BD9BB7 CRC64;

2.8%; Score 8; DB 4; Length 343;
rity 100.0%; Pred. No. 45;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLGLAL 67

Db 311 ALGLGLAL 318
|||||

RESULT 49

Q8TUU8

ID Q8TUU8 PRELIMINARY; PRT; 370 AA.

AC Q8TUU8;

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Permease subunit of a ABC-type transport system involved in

lipoprotein release.

GN MK1655

OS Methanopyrus kandleri.

OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyrac

OC Methanopyrus.

OX NCBI_TaxID=2320;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-AV19 / DSM 6324 / JCM 9639;

RA MEDLINE=21927647; PubMed=11930014;

RA Slesarev A.I., Mezheva V.V., Makarova K.S., Polushin N.N.,

RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,

RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,

RA Malykh A.G., Koonin E.V., Kozhavkin S.A.;

RT "The complete genome of hyperthermophile Methanopyrus kandleri AV.

and monophyly of archaeal methanogens.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).

DR EMBL; AB010455; AA002868.1; -.

DR GO; GO:0016020; C:membrane; IEA.

DR InterPro: IPR003838; DUF214.

DR Pfam; PF02687; FtsX; 1.

KW Complete proteome.

SQ SEQUENCE 370 AA; 39411 MW; B07662EALB5A644E CRC64;

Query Match 2.8%; Score 8; DB 17; Length 370;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 61 LGLGLALA 68

Db 336 LGLGLALA 343
|||||

RESULT 50

O53860

ID O53860 PRELIMINARY; PRT; 372 AA.

AC O53860;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Hypothetical protein cysM3.

GN CYSM3 OR RV0848 OR MT0043.41.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37Rv;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris

Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

[illegible]

REMBLrel. 13, Created)
 REMBLrel. 13, Last sequence update)
 REMBLrel. 25, Last annotation update)
 1.20 (At2g45000 protein).
 T2G45000.
 aliana (Mouse-ear cress).
 idiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Magnoliophyta; eudicotyledons; core eudicots; rosids;
 rassicales; Brassicaceae; Arabidopsis.
 2;
 N.A.
 H., Cheuk R., Kim C.J., Lim J., Meyers M.C., Banh J.,
 ninci P., Chang E., Dale J.M., Goldsmith A.D.,
 Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 dera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 wick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 u G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 DNA clones.";
 -2002) to the EMBL/GenBank/DBJ databases.
 N.A.
 S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 hen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 lin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 rusa M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 ull length cDNA Clones";
 -2002) to the EMBL/GenBank/DBJ databases.
 N.A.
 umbia;
 87; PubMed=10617197;
 son T.W., Bowman C.L., Barnstead M.E., Feidblym T.V.,
 tchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 hen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.R.,
 riera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 Preuss D., Nierman W.C., White O., Eisen J.A.,
 Fraser C.M., Venter J.C.;
 analysis of chromosome 2 of the plant Arabidopsis
 -768(1999).
 N.A.
 umbia;
 -2000) to the EMBL/GenBank/DBJ databases.
 ; AAL69462.1; -;
 ; AAL86303.1; -;
 ; AAD32835.1; -;
 84885.
 xotein.
 AA; 40584 MW; AF6C6B3BAC9BF69A CRC64;
 2.8%; Score 8; DB 10; Length 387;
 rity 100.0%; Pred. No. 51;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 EED 100
 ||||
 EED 377
 ELIMINARY; PRT; 397 AA.

DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein Atu3948.
 GN ATU3948 OR AGR_L_1808.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 Kutyavyn T., Levy R., Li M.-J., McClelland E., Palmeri A., Gordon
 Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon
 Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaci
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin
 RA Gurolo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin
 RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz E
 RA Planagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 Cielo C., Slater S;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE009325; AAL44750.1; ALT_INIT.
 DR EMBL; AE008289; AAK89478.1; -;
 DR PIR; AH3041; AH3041.
 DR PIR; D98244; D98244.
 DR InterPro; IPR001220; Lectin legB.
 DR InterPro; IPR001608; UPF0001.
 DR Pfam; PF01168; Ala racemase N; 1.
 DR PROSITE; PS00307; LECTIN LEGUME BETA; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 397 AA; 41708 MW; 700748E32A46AE86 CRC64;
 Query Match 2.8%; Score 8; DB 16; Length 397;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps
 QY 251 LALRPGSS 258
 Db. 341 LALRPGSS 348
 |||||
 |||||
 RESULT 56
 QYU9H4
 ID QYU9H4 PRELIMINARY; PRT; 431 AA.
 AC QYU9H4;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Possible bicarbonate transporter, ICT family.
 DE SYNW0284.
 GN SynW0284.
 OS Synchococcus sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
 OX NCBI_TaxID=84588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22825697; PubMed=12917641;
 RA Palenik B., Brahamsha B., Larimer F.W., Land M., Hauser L., Chain

Regala W., Allen E.E., McCarren J., Paulsen I.,
Partensky F., Webb E.A., Waterbury J.;
if a motile marine Synechococcus";
37-1042(2003).
9; CAB06799.1; --
eome.
1 AA; 46300 MW; 56295F913903DBAE CRC64;
2.8%; Score 8; DB 16; Length 431;
arity 100.0%; Pred. No. 56;
onservative 0; Mismatches 0; Indels 0; Gaps
GLLL 74
||||
GLLL 371
RELIMINARY; PRT; 435 AA.
TremBLrel. 20, Created)
TremBLrel. 20, Last sequence update)
TremBLrel. 25, Last annotation update)
smembrane protein.
.03756.
anacearum (Pseudomonas solanacearum).
lamid.
teobacteria; Betaproteobacteria; Burkholderiales;
aeae; Ralstonia.
15;
{ N.A.
10;
879; PubMed=11823852;
Genin S., Ariguenave F., Gouzy J., Mangenot S.,
lault A., Brottier P., Camus J.C., Cattolico L.,
Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
vie M., Moisan A., Robert C., Saurin W., Schlex T.,
heault P., Whalen M., Wincker P., Levy M.,
f., Boucher C.A.;
nce of the plant pathogen Ralstonia solanacearum";
17-502(2002).
9; CAD17762.1; --
1; C:extrachromosomal DNA; IEA.
4; F:ATP binding; IEA.
0; F:nucleoside-diphosphate kinase activity; IEA.
1; P:CTP biosynthesis; IEA.
3; P:GTP biosynthesis; IEA.
8; P:UTP biosynthesis; IEA.
.001564; NDK.
469; NDP_KINASES; 1.
lete proteome.
5 AA; 47048 MW; CCD859D9C54DB5A CRC64;
2.8%; Score 8; DB 16; Length 435;
arity 100.0%; Pred. No. 56;
onservative 0; Mismatches 0; Indels 0; Gaps
VLAL 224
||||
VLAL 365
RELIMINARY; PRT; 465 AA.
TremBLrel. 20, Created)
TremBLrel. 20, Last sequence update)
TremBLrel. 25, Last annotation update)
brane protein (Putative permease, major facilitator

CN	YEGB OR YFO2850 OR Y1393.
OS	Yersinia pestis.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Yersinia.
NCBI_TaxID=632;	
[1]_TaxID=632;	
RP	SEQUENCE FROM N.A.
RP	STRAIN-CO-92 / Biovar Orientalis;
RC	MEDLINE=21470413; PubMed=11586360;
RX	ParKhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.
RA	Praetice M.B., Sebailia M., James K.D., Churcher C., Mungall K.L.
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.N.
RA	Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA	Felwell T., Hanlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA	Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT	"Genome sequence of Yersinia pestis, the causative agent of plague
RL	Nature 413:523-527(2001).
[2]	
RP	SEQUENCE FROM N.A.
RP	STRAIN-KIMS / Biovar Mediaevalis;
RX	MEDLINE=22137863; PubMed=12142430;
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liu
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner
RA	Perry R.D.;
RT	"Genome sequence of Yersinia pestis KIM. ";
RL	J. Bacteriol. 184:4601-4611(2002).
DR	EMBL; AF414154; AAC92102.1; -
DR	EMBL; AF013741; AAC84955.1; -
DR	PIR; AC0347; AC0347.
DR	GO; GO:0016031; C:integral to membrane; IEA.
DR	GO; GO:0015520; P:tetracycline:hydrogen antiporter activity; IEA
DR	GO; GO:0005215; P:transporter activity; IEA.
DR	GO; GO:0015904; P:tetracycline transport; IEA.
DR	GO; GO:0006810; P:transport; IEA.
DR	InterPro; IPR007114; MFS.
DR	InterPro; IPR005828; Sub transporter.
DR	InterPro; IPR001411; TCR TetB.
DR	Pfam; PF00083; sugar tr_1.
DR	PRINTS; PR01036; TCR TetB.
DR	PROSITE; PS00850; MFS; 1.
DR	Hypothetical protein; Complete proteome.
KW	SEQUENCE 465 AA; 50176 MW; 0CC273F10B83F5ED CRC64;
QY	Query Match 2.8%; Score 8; DB 16; Length 465;
Db	Best Local Similarity 100.0%; Pred. No. 60;
	Matches 8; Conservative 0; Mismatches 0; Indels 0;
	72 LLLAVVSL 79
	335 LLLAVVSL 342
RESULT 59	
Q9RR18	PRELIMINARY; PRT; 471 AA.
ID	Q9RR18
AC	Q9RR18;
DT	01-MAY-2000 (TReMBLrel. 13, Created)
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE	Transport protein, putative.
GN	DR2502.
OC	Deinococcus radiodurans.
OC	Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC	Deinococcaceae; Deinococcus.
OC	NCBI_TaxID=1299;
RP	[1]_TaxID=1299;
RP	SEQUENCE FROM N.A.
RP	STRAIN-R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX	MEDLINE=20036896; PubMed=10557266;
RA	White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.

aft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
in H., Jiang L., Pamphile W., Crosby M., Shen M.,
Lam P., McDonald L., Utterback T., Zalewski C.,
Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
ce of the radioreistant bacterium Deinococcus
";
71-1577(1999).
; AAF12043.1; -.
75267.
-.
ome.
AA; 47974 MW; 96B2BEBF6E445D27 CRC64;

2.8%; Score 8; DB 16; Length 471;
rity 100.0%; Pred. No. 61;
nservative 0; Mismatches 0; Indels 0; Gaps 0;
HAL 67
|||||
HAL 373

ELIMINARY; PRT; 472 AA.

REMBLrel. 19, Created)
REMBLrel. 19, Last sequence update)
REMBLrel. 22, Last annotation update)
rotein FLJ31346.

Human).
azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
eria; Primates; Catarrhini; Hominidae; Homo.
6;

N.A.
mazaki M., Watanabe K., Kumagai A., Itakura S.,
ujimori Y., Kamiyama M., Sugiyama T., Irie R.,
O.H., Wakatsuki A., Ishii S., Yamamoto J., Isono Y.,
Saito K., Nishikawa T., Kimura K., Yamashita H.,
amuro Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
anehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
huzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,

NA sequencing project."
-2001) to the EMBL/GenBank/DBJ databases.

; BAB71043.1; -.
04299; MBOAT_fam.

protein; 1.

; AA; 52774 MW; EA721998043F9EBD CRC64;

2.8%; Score 8; DB 4; Length 472;
rity 100.0%; Pred. No. 61;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

HAL 67

|||||

HAL 447

ELIMINARY; PRT; 473 AA.

REMBLrel. 21, Created)
REMBLrel. 21, Last sequence update)
REMBLrel. 22, Last annotation update)
30589L02 gene.

(Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC023417; AAH23417.1; -.
DR MGD; MGI:1924832; 5730589L02Rik.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1
SQ SEQUENCE 473 AA; 53382 MW; DAALFEODA78013EA CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 65 LALACLGL 72
|||||
DB 436 LALACLGL 443

RESULT 62

Q9CY76

ID Q9CY76 PRELIMINARY; PRT; 473 AA.

AC Q9CY76;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE 5730589L02Rik protein.

GN 5730589L02Rik.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.

RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush

RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmit

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.

RA Hayashizaki Y.

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK019981; BAB31950.1; -.
DR MGD; MGI:1924832; 5730589L02Rik.

DR InterPro; IPR004299; MBOAT_fam.

DR Pfam; PF03062; MBOAT; 1.

SQ SEQUENCE 473 AA; 53504 MW; CE6F8E93C3D01C4F CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 65 LALACLGL 72
|||||
DB 436 LALACLGL 443


```

RE Q9LIW0 PRELIMINARY; PRT; 473 AA.
AC Q9LIW0;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE Similar to an Arabidopsis thaliana chromosome BAC genomic
DE sequence.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsing Y.C., Chow T., Chen C., Wu H., Chu M., Chao Y., Liu S.;
RT "Oryza sativa PAC P6699E04 genomics sequence, complete sequence."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001111; BAA90509.1; -.
DR Gramene; Q9LIW0; -.
SQ SEQUENCE 522 AA; 54697 MW; 21C6BAD2441B56BF CRC64;

Query Match 2.8%; Score 8; DB 10; Length 522;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 43 RRRGRGE 50
DB 415 RRRGRGE 422

RESULT 65
Q9LIW0 PRELIMINARY; PRT; 522 AA.
AC Q9LIW0;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE Similar to an Arabidopsis thaliana chromosome BAC genomic
DE sequence.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsing Y.C., Chow T., Chen C., Wu H., Chu M., Chao Y., Liu S.;
RT "Oryza sativa PAC P6699E04 genomics sequence, complete sequence."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001111; BAA90509.1; -.
DR Gramene; Q9LIW0; -.
SQ SEQUENCE 522 AA; 54697 MW; 21C6BAD2441B56BF CRC64;

Query Match 2.8%; Score 8; DB 10; Length 522;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 43 RRRGRGE 50
DB 415 RRRGRGE 422

RESULT 66
Q7WZ71 PRELIMINARY; PRT; 535 AA.
AC Q7WZ71;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Putative mannosyltransferase.
GN DBV20.
OS Nonomuraea sp. ATCC 39727.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Streptosporangiaceae; Nonomuraea.
OX NCBI TaxID=93944;
RN [1]
RP SEQUENCE FROM N.A.
RA Sosló M., Stinché S., Beltrametti F., Lazzarini A., Donadio S.;
RT "The gene cluster for the biosynthesis of the glycopeptide antib.
RT A40926 by Nonomuraea sp."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ561198; CAD91215.1; -.
KW Acyltransferase; Glycosyltransferase; Monooxygenase; Transferase
SQ SEQUENCE 535 AA; 57418 MW; 3C9059338B3308AC CRC64;

Query Match 2.8%; Score 8; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 122 RLVRPRRS 129
DB 480 RLVRPRRS 487

RESULT 67
Q8PMH8 PRELIMINARY; PRT; 537 AA.
AC Q8PMH8;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Oligopeptide transporter.

```

0. onopodis (pv. citri).
 eobacteria; Gammaproteobacteria; Xanthomonadales;
 ae; Xanthomonas.
 29;
 N.A.
 TCC 13902 / XV 101;
 45; PubMed12024217;
 ,; Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
 Cannavan F., Cardoso J., Chambergo F., Chapina L.P.,
 B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 reira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 F., Franco M.C., Greggio C.C., Gruber A.,
 Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 Santos M., Truffi D., Tsai S.M., White F.F.,
 Kitajima J.P.;
 The genomes of two Xanthomonas pathogens with differing
 ties.";
 1-463 (2002).
 ; AAM36320.1; -;
 ; C-membrane; IEA.
 ; F-transporter activity; IEA.
 ; Polypeptide transport; IEA.
 00109; PTR2.
 PTR2; 1.
 some.
 ; AA; 58369 MW; 7844COCOFEE8670 CRC64;
 2.8%; Score 8; DB 16; Length 537;
 rity 100.0%; Pred. No. 68;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 AVS 78
 ||||
 AVS 263
 ELIMINARY; PRT; 564 AA.
 REMBLrel. 10, Created)
 REMBLrel. 10, Last sequence update)
 REMBLrel. 25, Last annotation update)
 H5 (Fragment).
 A.
 negative-strand viruses; Orthomyxoviridae;
 ruses.
 911;
 N.A.
 Potadam/2216-4/84;
 02; PubMed9882316;
 ; Zhou N., Kawaoka Y., Webster R.;
 glycoproteins of H5 influenza viruses isolated from
 ns, and wild aquatic birds have distinguishable
 146-1155(1999).
 HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
 TORS AND FOR INITIATING INFECTION (BY SIMILARITY).
 HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
 (A2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
 ; AAD13573.1; -;
 1HTM.

DR GO: GO:0019031; C:viral envelope; IEA.
 DR InterPro; IPR008980; Capsid hemag.
 DR InterPro; IPR001364; Hemagglutn.
 DR Pfam; PF00509; Hemagglutinin; 1.
 DR PRINTS; PR00329; HEMAGGLUTN12.
 DR ProDom; PD000225; Hemagglutn; 1.
 KW Envelope protein; Glycoprotein; Hemagglutinin.
 FT NON TER 564 564
 SQ SEQUENCE 564 AA; 63562 MW; B317179A7F3E6F98 CRC64;
 Query Match 2.8%; Score 8; DB 12; Length 564;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G
 QY 72 LLLAVVSL 79
 |||||
 Db 6 LLLAVVSL 13
 RESULT 69
 Q8JN92 PRELIMINARY; PRT; 568 AA.
 AC Q8JN92;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hemagglutinin H5.
 GN H5.
 OS Influenza A virus (A/Goose/Hong Kong/3014.5/2000(H5N1)).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Influenza A viruses; Influenzavirus A.
 OX NCBI_TaxID=186167;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A/Goose/Hong Kong/3014.5/2000;
 RX MEDLINE-22016166; PubMed-12021367;
 RA Tumpey T.M., Suarez D.L., Perkins L.E.L., Senne D.A., Lee J.G.,
 Lee Y.J., Mo I.P., Sung H.W., Swayne D.E.;
 RT "Characterization of a Highly Pathogenic H5N1 Avian Influenza A V:
 Isolated from Duck Meat";
 RL J. Virol. 76:6344-6355(2002).
 CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRU:
 CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
 CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CH:
 (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
 DR EMBL; AY075030; AA175843.1; -;
 DR GO: GO:0019031; C:viral envelope; IEA.
 DR InterPro; IPR008980; Capsid hemag.
 DR InterPro; IPR001364; Hemagglutn.
 DR Pfam; PF00509; Hemagglutinin; 1.
 DR PRINTS; PR00329; HEMAGGLUTN12.
 DR ProDom; PD000225; Hemagglutn; 1.
 KW Envelope protein; Glycoprotein; Hemagglutinin.
 SQ SEQUENCE 568 AA; 64244 MW; E0D741A75C6E76FC CRC64;
 Query Match 2.8%; Score 8; DB 12; Length 568;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G
 QY 72 LLLAVVSL 79
 |||||
 Db 6 LLLAVVSL 13
 RESULT 70
 Q8QPL0 PRELIMINARY; PRT; 568 AA.
 ID Q8QPL0;
 AC Q8QPL0;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hemagglutinin (Fragment).

```

RT      host specificities. ";
RRL     Nature 417:459-463(2002).
R      EMBL; AE012240; AAM40703.1; -.
GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005215; F:transporter activity; IEA.
DR      GO; GO:0006857; P:oligopeptide transport; IEA.
DR      InterPro; IPR00103; PTR2.
DR      Pfam; PF00854; PTR2; 1.
DR      PROSITE; PS01022; PTR2_1; 1.
KW      Complete proteome.
SQ      SEQUENCE 620 AA; 67314 MW; E9904BFF039B6AEC CRC64;

Query Match      2.8%; Score 8; DB 16; Length 620;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY      65 LALACLGL 72
DB      349 LALACLGL 356
      |||||
      |||||

RESULT 72
Q9N8H2      PRELIMINARY; PRT; 656 AA.
ID Q9N8H2
AC Q9N8H2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN TB927.1.3840.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypano-
OX NCBI_TaxID=5691;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TREU927;
RR Hall N., Beiraman M., Lennard N.J., Harris B.R., Gerrard C.S.,
RA Atkin R.J., Barron A.J., Bart-Delabesse E.N., Bowman S.,
RA Bray-Allen S.P., Bringaud F., Clark L.N., Corton C.H., Cronin A.,
RA Davies R., Doggett J., Fraser A., Gruter E., Hall S., Harper D.A.,
RA Hertz-Powler C., Kay M.P., Leech V., Mayes R., Price C., Quail M.R.,
RA Rabinowitsch E., Rutherford K., Sasse J., Sharp S., Showkneen R.,
RA Gull K., Barrell B.G., Melville S.E.;
RR "The sequence and analysis of the highly polymorphic chromosome 6:
RA the African trypanosome, Trypanosoma brucei."
RRL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL929607; CAB95571.1; -.
KW Hypothetical protein.
SQ      SEQUENCE 656 AA; 72138 MW; CBAC892D25937FAD CRC64;

Query Match      2.8%; Score 8; DB 5; Length 656;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY      93 EELVAED 100
DB      454 EELVAED 461
      |||||
      |||||

RESULT 73
Q9N8B      PRELIMINARY; PRT; 1523 AA.
ID Q9N8B
AC Q9N8B;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN TB927.1.1600.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypano-
OX NCBI_TaxID=5691;
RX [1]
RP SEQUENCE FROM N.A.

```

7;
 iman M., Lennard N.J., Harris B.R., Gerrard C.S.,
 aron A.J., Bart-Delabesse E.N., Bowman S.,
 P., Bringuand F., Clark L.N., Corton C.H., Cronin A.,
 Jiggett J., Fraser A., Gruter E., Hall S., Harper D.A.,
 2, Kay M.P., Leech V., Mayes R., Price C., Quail M.A.,
 E., Rutherford K., Sasse J., Sharp S., Showkeen R.,
 ell B.G., Melville S.E.;
 and analysis of the highly polymorphic chromosome I of
 trypanosome, *Trypanosoma brucei*.";
 2-2002) to the EMBL/GenBank/DBJ databases.
 1; CAB95435.1; -
 001656; UPF0024.
 ; UPF0024; 1.
 protein.
 23 AA; 168322 MW; 077BDC751CDD1E5A CRC64;
 2.8%; Score 8; DB 5; Length 1523;
 urity 100.0%; Pred.No. 1.8e+02;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 .ALR 225
 ||||
 .ALR 389
 ELIMINARY; PRT; 32 AA.
 REMBLrel. 17, Created)
 REMBLrel. 17, Last sequence update)
 REMBLrel. 19, Last annotation update)
 ulin precursor (Fragment).
 ULIN.
 Goat).
 azoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 eria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 nae; Capra.
 5;
 N.A.
 illa F., Graziano M.;
 phism of goat beta-lactoglobulin proximal promoter
 (-2001) to the EMBL/GenBank/DBJ databases.
 ; CAC27455.1; -.
 1 18 POTENTIAL.
 2 32
 AA; 3372 MW; 0C56BD579B3DC190 CRC64;
 2.5%; Score 7; DB 6; Length 32;
 urity 100.0%; Pred.No. 52;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 AC 69
 ||||
 AC 14
 ELIMINARY; PRT; 35 AA.
 REMBLrel. 10, Created)
 REMBLrel. 10, Last sequence update)
 REMBLrel. 25, Last annotation update)
 rotein (Fragment).
 homatis.
 mydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 ;

06:25:19 2004

us-09-245-198a-4.oligo.rag

2.5 66 7 ADC97100
2.5 69 4 ABB65500

ALIGNMENTS

idard; protein; 284 AA.

(first entry)

tumour necrosis factor related ligand (TRELL).

r necrosis factor related ligand; tnfr; treatment; cancer;
isease; immune system; stimulation; suppression;
ion.

97WO-US013945.

96US-0023541P.

96US-0028515P.

97US-0040820P.

EN INC.

GENEVA FACULTY MEDICINE.

he Y, Browning JL;

5619/13.

9600.

sis factor related ligand - useful for, e.g. treating cancer,
isease and immune responses to tissue grafts.

je 50-51; 69pp; English.

is that of human tumour necrosis factor related ligand
LL or active fragments can be included with a carrier in
al compositions to treat cancer, autoimmune diseases or
ases to tissue grafts, or to stimulate or suppress the immune
s useful to screen for TRELL receptors, by labelling with a
abel and screening compositions for binding. Agents
with TRELL-receptor binding can also be screened for, can
istered, optionally with interferon- gamma, to induce cell
at, suppress or alter immune responses (especially involving
arcinoma cells) involving a signal pathway between TRELL and
. It's coding sequence can be used in gene therapy for TRELL-
rders in mammals (especially humans), e.g. tumours,
nd inflammatory diseases or inherited genetic disorders, by
into cells, and expressing, therapeutically effective amounts
e.g. a virus comprising a gene encoding TRELL. It may also
the preparation of prepare probes for screening
hetic DNAs for TRELL-encoding sequences and for antisense

AA;

100.0%; Score 284; DB 2; Length 284;

larity 100.0%; Pred. No. 2e-252;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LDFEISARRLPRLPSLGRDGGAVRQAQPPAPMAARRSQRRRRGGPGTALLVPLA 60

Db 1 MSLLDFEISARRLPRLPSLGRDGGAVRQAQPPAPMAARRSQRRRRGGPGTAL
QY 61 LGGLGALACLGLLAVVSLGSRASLSAQEPAQOEELVAEEDQDPSELNPOTESQD
Db 61 LGGLGALACLGLLAVVSLGSRASLSAQEPAQOEELVAEEDQDPSELNPOTESQD
QY 121 NLRVPRRSAPKGRKTRARRAJAAHYEVHPRPGDGAQAGVDTVSGWEARINS
Db 121 NLRVPRRSAPKGRKTRARRAJAAHYEVHPRPGDGAQAGVDTVSGWEARINS
QY 181 YNRQIGEFIVTRAGLYLYCYQVHFDEGKAVYKLDLLVDGVLALRCLEFSATAA
Db 181 YNRQIGEFIVTRAGLYLYCYQVHFDEGKAVYKLDLLVDGVLALRCLEFSATAA
QY 241 QLRLQVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYFGLFQVH 284
Db 241 QLRLQVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYFGLFQVH 284

RESULT 2

AAY09369

ID AAY09369 standard; protein; 249 AA.

XX AC AAY09369;

XX DT 15-JUL-1999 (first entry)

XX DE Human tumour necrosis factor Apo-3 ligand protein sequence.

XX KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apopto

XX KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;

XX OS Homo sapiens.

XX DN WO919490-A1.

XX PD 22-APR-1999.

XX PF 09-OCT-1998; 98WO-US021407.

XX PR 10-OCT-1997; 97US-0062037P.

XX PR 17-DEC-1997; 97US-0069862P.

XX PA (GETH) GENENTECH INC.

XX PI Ashkenazi AJ, Marsters SA, Pitti R;

XX DR WPI; 1999-287982/24.

XX DR N-PSDB; AAX56000.

XX PT New human Apo3- ligand (a tumor necrosis factor) homologue.

XX PS Claim 1; Fig 1; 74pp; English.

XX CC The present sequence represents a human tumour necrosis factor (lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has cytostatic activity. Apo-3 ligand can be used to induce apoptosis in mammalian cancer cells, to induce NF-kappaB-dependent transcript to induce JNK/SAPK-dependent responses in mammalian cells

XX SQ Sequence 249 AA;

Query Match 87.7%; Score 249; DB 2; Length 249;

Best Local Similarity 100.0%; Pred. No. 2.6e-220;

Matches 249; Conservative 0; Mismatches 0; Indels 0;

QY 36 MAARSQRRRRGGPGTALLVPLALGLGALACLGLLAVVSLGSRASLSAQEP.

Db 1 MAARSQRRRRGGPGTALLVPLALGLGALACLGLLAVVSLGSRASLSAQEP.

QY 96 VAEEDQDPSELNPOTESQDPAPFLNRLVPRRSAPKGRKTRARRAJAAHYEVH.

Db 61 VAEEDQDPSELNPOTESQDPAPFLNRLVPRRSAPKGRKTRARRAJAAHYEVH.

WDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYLYCQVHFDEGKAVYLKLD 215
 |||||
 WDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYLYCQVHFDEGKAVYLKLD 180
 |||||
 WDLALCLFEFSATAASSLGPQLRLCQVSGLLALRPGSSLRINTLPWAHLKAAPFL 275
 |||||
 WDLALCLFEFSATAASSLGPQLRLCQVSGLLALRPGSSLRINTLPWAHLKAAPFL 240
 |||||
 FQVH 284
 |||||
 FQVH 249

hard; protein; 249 AA.

(first entry)

antitumour protein.

antitumour; tumour; therapy; cytostatic; breast cancer;
 ; renal cancer; colorectal cancer; uterine cancer;
 ; lung cancer; bladder cancer;
 is system cancer; melanoma; leukaemia; neoplasm.

Location/Qualifiers

1. .40
 /label= signal_peptide
 10. .14
 /note= "amidation"
 24. .35
 /note= "prokaryotic membrane lipoprotein lipid"
 27. .33
 /note= "N-myristoylation"
 29. .35
 /note= "N-myristoylation"
 36. .42
 /note= "N-myristoylation"
 41. .249
 /label= PRO207
 45. .51
 /note= "N-myristoylation"
 97. .101
 /note= "amidation"
 118. .124
 /note= "N-myristoylation"
 121. .127
 /note= "N-myristoylation"
 125. .131
 /note= "N-myristoylation"
 128. .134
 /note= "N-myristoylation"
 139. .143
 /note= "Asn is N-glycosylated"

2.

99WO-US028565.
 98US-0113296P.
 99WO-US005028.
 99US-0130232P.
 99US-0131445P.
 99US-0134287P.
 99US-0144758P.

PR 26-JUL-1999; 99US-0145698P.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 XX
 PA (GEHT) GENENTECH INC.
 XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
 PI Napier MA, Pitti RM, Wood WI;
 XX WPI; 2000-442668/38.
 DR N-PSDB; AAA49717.
 XX
 PT Novel composition to inhibit neoplastic cell growth or for treati
 in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
 PT PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866.
 XX
 PS Claim 19; Fig 4; 172pp; English.
 XX
 CC The present sequence is that of human antitumour protein PRO207,
 deduced from a foetal kidney cDNA clone (see AAA49717). PRO207 sh
 CC amino acid sequence identity to tumour necrosis factor family mem
 CC especially human lymphotoxin-beta (23.4%) and human CD40 ligand (
 CC Mol wt. is 27,216. A claimed method for inhibiting the growth of
 CC cell comprises exposing the tumor cell to PRO179, PRO207, PRO320,
 CC PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or
 CC (see AAY95337-49), their agonists or chimeric polypeptides incorp
 CC them. The tumour is especially a cancer selected from breast, ova
 CC renal, colorectal, uterine, prostate, lung, bladder and central n
 CC system cancer, melanoma and leukaemia. Methods for the recombinan
 CC expression of the antitumour proteins are also provided
 XX
 SQ Sequence 249 AA;

Query Match 87.7%; Score 249; DB 3; Length 249;
 Best Local Similarity 100.0%; Pred. No. 2.6e-220;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; G

QY 36 MAARRSQRRGRGEGPTALLVPLALGLGLALACGLLLAVWSLGSRLSLAQEPA
 |||||
 DB 1 MAARRSQRRGRGEGPTALLVPLALGLGLALACGLLLAVWSLGSRLSLAQEPA
 |||||
 QY 96 VAEDQDPSELNPQTESQDPAPFLNLRVPRSPKGRKTRARRAAAHYEVHPR
 |||||
 DB 61 VAEDQDPSELNPQTESQDPAPFLNLRVPRSPKGRKTRARRAAAHYEVHPR
 |||||
 QY 156 GAQAGVDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYLYCQVHFDEGKAVY
 |||||
 DB 121 GAQAGVDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYLYCQVHFDEGKAVY
 |||||
 QY 216 LLVDGVLALRCLFEFSATAASSLGPQLRLCQVSGLLALRPGSSLRINTLPWAHLKA
 |||||
 DB 181 LLVDGVLALRCLFEFSATAASSLGPQLRLCQVSGLLALRPGSSLRINTLPWAHLKA
 |||||
 QY 276 TYEGLFQVH 284
 |||||
 DB 241 TYEGLFQVH 249

RESULT 4

AAB07526
 ID AAB07526 standard; protein; 249 AA.
 XX
 AC AAB07526;
 XX
 DT 20-OCT-2000 (first entry)
 XX

DE Amino acid sequence of a soluble recombinant human TWEAK protein.
 XX
 XX TWEAK protein; immunological disorder; immune response; inflammat
 KW TWEAK blocking agent; autoimmune disease; organ transplant reject
 KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock;
 XX
 OS Homo sapiens.

QVLAIRCLEEFSAATAASSLGPQLRLCOVSLALRPQSSLRIRTLPLWAHLK 240

LFQVH 284

LFQVH 249

dard; protein; 249 AA.

(first entry)

rotein.

tumour necrosis factor; ligand; cytostatic;
or; osteopathic.

A2.

2002WO-US023782.

2001US-0307838P.

GENOME SCI INC.

Rosen CA;

659/40.

901.

timeric complex having a first polypeptide member of the
s factor (TNF) ligand family, and a second different member
family, useful for treating cancer, osteoporosis or an
sease.

age 368-369; 388pp; English.

equene is the protein sequence for human TWEAK protein. The
ates to compositions comprising heterotrimeric complexes of
is factor (TNF) ligand family members, and their use in the
vention and treatment of disease. In one embodiment, the
c complex comprises full-length or extracellular portions of
1-length or extracellular portions of other TNF ligand
s, preferably VEGI or VEGI-SV. The heterotrimeric complexes
ion are useful for treating an autoimmune disease, cancer or
and particularly for inhibiting cancer cell proliferation,
cell proliferation, or inducing apoptosis of T cells

AA;

87.7%; Score 249; DB 6; Length 249;

arity 100.0%; Pred. No. 2.6e-220;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

RSQRRRGRRGPGTALLVPLALGLALACLGILLAVVSLGSRASLSAQEPQDEL 95

RSQRRRGRRGPGTALLVPLALGLALACLGILLAVVSLGSRASLSAQEPQDEL 60

QDPSELNPQTESQDPAPFLNLRVRRSAPKGRKTRARRAJAAHYEHVHPRPGQD 155

QDPSELNPQTESQDPAPFLNLRVRRSAPKGRKTRARRAJAAHYEHVHPRPGQD 120

3VDGTSGWEARINSSSPRYNRQIGEFIVTRAGLYLYLCOVHFDGKAVYIKLD 215

3VDGTSGWEARINSSSPRYNRQIGEFIVTRAGLYLYLCOVHFDGKAVYIKLD 180

QY 216 LLVDGVLAIRCLEEFSAATAASSLGPQLRLCOVSLALRPQSSLRIRTLPLWAHLK

Db 181 LLVDGVLAIRCLEEFSAATAASSLGPQLRLCOVSLALRPQSSLRIRTLPLWAHLK

QY 276 TYFGLFQVH 284

Db 241 TYFGLFQVH 249

RESULT 7

ADC35206

ID ADC35206 standard; protein; 249 AA.

XX

AC ADC35206;

XX

DT 18-DEC-2003 (first entry)

XX

Human TNF ligand family member #12.

XX

KW human; tumour necrosis factor; TNF ligand; endokine alpha;
excessive bone resorption disorder; osteoporosis; Paget's disease;

KW arterial calcification.

XX

OS Homo sapiens.

XX

US2003100074-A1.

XX

PD 29-MAY-2003.

XX

PF 15-AUG-2002; 2002US-00218547.

XX

PR 16-AUG-2001; 2001US-0312542P.

XX

PR 30-OCT-2001; 2001US-0330761P.

XX

PA (YUGG/) YU G.

XX

PA (NIJJ/) NI J.

XX

PA (ROSE/) ROSEN C A.

XX

PA (NARD/) NARDELLI B.

XX

PI Yu G, Ni J, Rosen CA, Nardelli B;

XX

WPI; 2003-696072/66.

DR

N-PSDB; ADC35205.

XX

PT New Endokine alpha gene useful for preparing a composition for tr
disease associated with excessive or insufficient bone resorptior

PT

PT osteoporosis, Paget's disease or arterial calcification.

XX

PS Disclosure; SEQ ID NO 24; 145pp; English.

XX

CC The invention relates to an isolated nucleic acid molecule encodi
tumour necrosis factor family ligand. A composition comprising th

CC

CC isolated antibody or its fragment is used for treating an individ
need of decreased level of endokine alpha activity. The endokine

CC

CC polypeptide present in a heterotrimeric complex is used for treat
individual having a disorder associated with excessive bone resor

CC

CC e.g. osteoporosis, Paget's disease or arterial calcification. The
individual having a disorder associated with insufficient bone re

CC

CC comprises administering an endokine alpha antagonist, which is th
antibody that binds specifically to endokine alpha polypeptide. T

CC

CC present sequence represents the amino acid sequence of a tumour n
factor family ligand.

XX Sequence 249 AA;

Query Match

Best Local Similarity 87.7%; Score 249; DB 7; Length 249;

Matches 249; Conservative 0; Mismatches 0; Indels 0; G

QY

36 MAARRSQRRRGRGPGTALLVPLALGLALACLGILLAVVSLGSRASLSAQEPA

Db

1 MAARRSQRRRGRGPGTALLVPLALGLALACLGILLAVVSLGSRASLSAQEPA

06:25:19 2004

us-09-245-198a-4.oligo.rag

EDQPSLNQTESQDPAPFLNRLVPRRSAPKGRKTRARRAJAAHYEVHPRPGQD 155
|||||
EDQPSLNQTESQDPAPFLNRLVPRRSAPKGRKTRARRAJAAHYEVHPRPGQD 120
|||||
AGVDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYYCQVHFDEGKAVYLKLD 215
AGVDGTVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYYCQVHFDEGKAVYLKLD 180
DGVLAIRCLBEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPAWHLKAAPFL 275
DGVLAIRCLBEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPAWHLKAAPFL 240
3LFQVH 284
3LFQVH 249
ndard; protein; 249 AA.

(first entry)
endothelium proliferative agent protein.
lium proliferative agent; TREPA; wound healing; cancer;
ing; vascularisation; apoptosis; autoimmune; birth control.

98WO-US002859.
97US-00798692.
98US-00021706.
IT LAB.

7255/38.
7613.
cleic acid encoding TREPA - useful for diagnosis and
autoimmune disease, tumours and inflammation.
3e 123-4; 142pp; English.
ted endothelium proliferative agent (TREPA), or its
agonists, are used to treat a deficit of TREPA, e.g. to
a healing or tissue grafting, by promoting vascularisation,
ce apoptosis for treating cancer and eliminating autoreactive
an adjunct to cancer chemotherapy or antiviral treatment.
es can also be used to target cytotoxic agents or for
lation of the corresponding receptor, the nucleic acid for
used to transform tumour cells to render them more
o TREPA and to screen for TREPA mimics. Ribozymes, antisense
dies or peptides, are used to treat TREPA-associated
g. tumours and metastases (by inhibiting vascularisation),
or a wide range of autoimmune conditions, conditions
normal stimulation of epithelial cells (e.g.
sis), for birth control (inhibiting ovulation and placental
r other angiogenic conditions (e.g. ulcers)
AA;
84.9%; Score 241; DB 2; Length 249;
larity 100.0%; Pred. No. 5.9e-213;
9

Matches 241; Conservative 0; Mismatches 0; Indels 0;
QY 44 RRRRGCEPTALLVPLALGLGLALACLGLLAVVSLGSRASLSAQEPAGEELVAE
Db 9 RRRRGCEPTALLVPLALGLGLALACLGLLAVVSLGSRASLSAQEPAGEELVAE
QY 104 SELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAJAAHYEVHPRPGQDGAQ
Db 69 SELNPQTESQDPAPFLNRLVPRRSAPKGRKTRARRAJAAHYEVHPRPGQDGAQ
QY 164 TVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYYCQVHFDEGKAVYLKLDLLV
Db 129 TVSGWEEARINSSPLRYNRQIGEFIVTRAGLYLYYCQVHFDEGKAVYLKLDLLV
QY 224 LRCLEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYF
Db 189 LRCLEFSATAASSLGPQLRLCQVSGLLALRPGSSLRIRTLPAWHLKAAPFLTYF
QY 284 H 284
Db 249 H 249

RESULT 9
AAE00891
ID AAE00891 standard; protein; 249 AA.
XX
AC AAE00891;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent).
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulneryary.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 98..249
FT /label= Extracellular_domain
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-00105343.
XX
PR 12-FEB-1997; 97US-00798692.
PR 10-FEB-1998; 98US-00021706.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI; 2001-280760/29.
DR N-PSDB; AAD04350.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting w
PT healing, by administering soluble fragment of extracellular doma
PT tumor necrosis factor related endothelium proliferative agent pr
PS Claim 1; Col 75-76; 53pp; English.
XX
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family mole
CC designated as TREPA (TNF related endothelium proliferative agent
CC Soluble biologically active TREPA are used to treat TREPA-associ
CC diseases, tumours or metastases. TREPA is used for inducing angi
CC in human for promoting wound healing and for vascularising graft
CC for successful grafting and to promote tissue grafts. The presen
CC acid sequence is clone ID #690050 human TREPA

XX
 CC The TNF-related endothelium proliferative agent (TREPA), or its
 CC activators or agonists, are used to treat a deficit of TREPA, e.
 CC promote wound healing or tissue grafting, by promoting vasculari-
 CC also to induce apoptosis for treating cancer and eliminating aut
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treat
 CC TREPA peptides can also be used to target cytotoxic agents or fo
 CC affinity isolation of the corresponding receptor, the nucleic ac
 CC which can be used to transform tumour cells to render them more
 CC responsive to TREPA and to screen for TREPA mimics. Ribozymes, a
 CC RNA, antibodies or peptides, are used to treat TREPA-associated
 CC diseases, e.g. tumours and metastases (by inhibiting vascularisa
 CC inflammation or a wide range of autoimmune conditions, condition
 CC involving abnormal stimulation of epithelial cells (e.g.
 CC atherosclerosis), for birth control (inhibiting ovulation and pl
 CC formation) or other angiogenic conditions (e.g. ulcers)
 XX
 SQ Sequence 189 AA;
 Query Match 50.4%; Score 143; DB 2; Length 189;
 Best Local Similarity 100.0%; Pred. No. 5.7e-123;
 Matches 143; Conservative 0; Mismatches 0; Indels 0;
 QY 142 IAAHYEVHPRPGDGAQAGVDGTGSGWEARINSSSPLRYNROI GEFIVTRAGLY
 Db 47 IAAHYEVHPRPGDGAQAGVDGTGSGWEARINSSSPLRYNROI GEFIVTRAGLY
 QY 202 VHFDEGKAVYKLDLLVDGVLALRCLEEFSAATAASSLGPGQLRCQVSGLLALRPG
 Db 107 VHFDEGKAVYKLDLLVDGVLALRCLEEFSAATAASSLGPGQLRCQVSGLLALRPG
 QY 262 RTLPWAHLKAAPFLTYFGLFQVH 284
 Db 167 RTLPWAHLKAAPFLTYFGLFQVH 189
 RESULT 13
 AAE00892
 ID AAE00892 standard; protein; 189 AA.
 AC AAE00892;
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE Human UL4flag TREPA soluble construct.
 XX
 KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
 KW TREPA; TNF related endothelium proliferative agent; metastasis;
 XX vulnery; HUVEC; human umbilical vein endothelial cell; UL4flag
 OS Homo sapiens.
 XX
 PN US6207642-B1.
 XX
 PD 27-MAR-2001.
 XX
 PF 26-JUN-1998; 98US-00105343.
 XX
 PR 12-FEB-1997; 97US-00798692.
 PR 10-FEB-1998; 98US-00021706.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Wiley SR;
 XX
 DR WPI; 2001-280760/29.
 XX
 PT Inducing angiogenesis in mammal at desired sites for promoting w
 PT healing, by administering soluble fragment of extracellular doma
 PT tumor necrosis factor related endothelium proliferative agent pr
 XX
 PS Example 2; Col 75-78; 53pp; English.
 XX

XX
 CC The TNF-related endothelium proliferative agent (TREPA), or its
 CC activators or agonists, are used to treat a deficit of TREPA, e.
 CC promote wound healing or tissue grafting, by promoting vasculari-
 CC also to induce apoptosis for treating cancer and eliminating aut
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treat
 CC TREPA peptides can also be used to target cytotoxic agents or fo
 CC affinity isolation of the corresponding receptor, the nucleic ac
 CC which can be used to transform tumour cells to render them more
 CC responsive to TREPA and to screen for TREPA mimics. Ribozymes, a
 CC RNA, antibodies or peptides, are used to treat TREPA-associated
 CC diseases, e.g. tumours and metastases (by inhibiting vascularisa
 CC inflammation or a wide range of autoimmune conditions, condition
 CC involving abnormal stimulation of epithelial cells (e.g.
 CC atherosclerosis), for birth control (inhibiting ovulation and pl
 CC formation) or other angiogenic conditions (e.g. ulcers)
 XX
 SQ Sequence 189 AA;
 Query Match 50.4%; Score 143; DB 2; Length 189;
 Best Local Similarity 100.0%; Pred. No. 5.7e-123;
 Matches 143; Conservative 0; Mismatches 0; Indels 0;
 QY 142 IAAHYEVHPRPGDGAQAGVDGTGSGWEARINSSSPLRYNROI GEFIVTRAGLY
 Db 47 IAAHYEVHPRPGDGAQAGVDGTGSGWEARINSSSPLRYNROI GEFIVTRAGLY
 QY 202 VHFDEGKAVYKLDLLVDGVLALRCLEEFSAATAASSLGPGQLRCQVSGLLALRPG
 Db 107 VHFDEGKAVYKLDLLVDGVLALRCLEEFSAATAASSLGPGQLRCQVSGLLALRPG
 QY 262 RTLPWAHLKAAPFLTYFGLFQVH 284
 Db 167 RTLPWAHLKAAPFLTYFGLFQVH 189
 RESULT 13
 AAE00892
 ID AAE00892 standard; protein; 189 AA.
 AC AAE00892;
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE Human UL4flag TREPA soluble construct.
 XX
 KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
 KW TREPA; TNF related endothelium proliferative agent; metastasis;
 XX vulnery; HUVEC; human umbilical vein endothelial cell; UL4flag
 OS Homo sapiens.
 XX
 PN US6207642-B1.
 XX
 PD 27-MAR-2001.
 XX
 PF 26-JUN-1998; 98US-00105343.
 XX
 PR 12-FEB-1997; 97US-00798692.
 PR 10-FEB-1998; 98US-00021706.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Wiley SR;
 XX
 DR WPI; 2001-280760/29.
 XX
 PT Inducing angiogenesis in mammal at desired sites for promoting w
 PT healing, by administering soluble fragment of extracellular doma
 PT tumor necrosis factor related endothelium proliferative agent pr
 XX
 PS Example 2; Col 75-78; 53pp; English.
 XX

vention relates to extracellular signal molecules, members of tumour necrosis factor (TNF) family molecules (TREPA (TNF related endothelium proliferative agent)). Gally active TREPA are used to treat TREPA-associated tumors or metastases. TREPA is used for inducing angiogenesis promoting wound healing and for vascularising grafted tissue and grafting and to promote tissue grafts. The present amino acid sequence is human U4fiag TREPA soluble construct. This sequence is biologically active molecule is capable of inducing HUVEC (human umbilical vein endothelial cells) cells

AA;

50.4%; Score 143; DB 4; Length 189;
arity 100.0%; Pred. No. 5.7e-123;
conservative 0; Mismatches 0; Indels 0; Gaps 0;

YEVHPRGQDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYYCQ 201
|||||
YEVHPRGQDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYYCQ 106

EGKAVYLKDLVDGLALRCLEEFSAATAASSLGPQLRLCQVSGLLALRPGSSLRI 261
|||||
EGKAVYLKDLVDGLALRCLEEFSAATAASSLGPQLRLCQVSGLLALRPGSSLRI 166

WAHLKAAPFLTYFGLFQVH 284
|||||
WAHLKAAPFLTYFGLFQVH 189

gard; protein; 208 AA.

(first entry)

rotein.

is factor receptor; signal transducer molecule; TNF; APO4; abnormality; gestational abnormality; prostate cancer; PO3; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease; main; immunogen; antibody preparation; breast carcinoma; nan.

98WO-US018393.

97US-00924634.

WASHINGTON.

191/17.
124.

rosis Factor family receptor polypeptides and ligands - agnosis and treatment of prostate cancer and developmental abnormalities.

13A; 156pp; English.

a describes isolated Tumor Necrosis Factor (TNF) family peptides: APO4, APO6, APO8 and APO9 or their active fragments. APO4 is useful for diagnosing prostate cancer by

CC determining levels of APO4 in an individual. Prostate cancer can
CC treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying select-
CC binding agents, useful in diagnosis/treatment of disease by bind-
CC agents to the polypeptide/active fragment which is extracellular,
CC expressed on the cell surface. The binding is preferably performed
CC vivo. APO4 polypeptides/ active fragments are also useful for sci-
CC for agonists and antagonists by binding and observing the change
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/
CC fragments and APO4 signal transducer molecules that specifically
CC with a cytoplasmic domain of APO4 and detecting a change in level
CC activity. The method is performed in vivo or in vitro. APO4 is al-
CC are all useful as immunogens for preparing antibodies. APO4 is al-
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cel-
CC MCF-7, and induced apoptosis

XX Sequence 208 AA;

Query Match 37.7%; Score 107; DB 2; Length 208;
Best Local Similarity 99.5%; Pred. No. 7.8e-90;
Matches 207; Conservative 0; Mismatches 1; Indels 0; C

Qy 77 VSLGSRASLSAQEPAQELVAEEDQDPSELNPQTEESQDPAPFLNRLVRRPSAPP
|||||
Db 1 VSLGSRASLSAQEPAQELVAEEDQDPSELNPQTEESQDPAPFLNRLVRRPSAPP

Qy 137 RARRAIAAHVEHPRPGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTR
|||||
Db 61 RARRAIAAHVEHPRPGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTR

Qy 197 YLYCOVHFDEGKAVYLKDLVDGLALRCLEEFSAATAASSLGPQLRLCQVSGLLP
|||||
Db 121 YLYCOVHFDEGKAVYLKDLVDGLALRCLEEFSAATAASSLGPQLRLCQVSGLLP

Qy 257 SSLRIRTLTPWAHLKAAPFLTYFGLFQVH 284

Db 181 SSLRIRTLTPWAHLKAAPFLTYFGLFQVH 208

RESULT 15

AAW93591

ID AAW93591 standard; protein; 211 AA.

XX AC AAW93591;

XX DT 18-JUN-1999 (first entry)

XX DE Mouse TNRL3 protein.

XX DE Tumour necrosis factor receptor; signal transducer molecule; TNF;
XX KW developmental abnormality; gestational abnormality; prostate ca-
XX KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy;
XX KW cytoplasmic domain; immunogen; antibody preparation; breast carci-
XX KW apoptosis; mouse.

OS Mus sp.

XX PN WO9911791-A2.

XX PD 11-MAR-1999.

XX PF 04-SEP-1998; 98WO-US018393.

XX PR 05-SEP-1997; 97US-00924634.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Chaudhary PM;

XX WK 1999-205191/17.

DR N-PSDB; AAX23425.

rosis Factor family receptor polypeptides and ligands -
agnosis and treatment of prostate cancer and developmental
al abnormalities.

g 13B; 156pp; English.

on describes isolated Tumor Necrosis Factor (TNF) family
ypeptides: APO4, APO6, APO8 and APO9 or their active
nd isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
ragments. APO4 is useful for diagnosing prostate cancer by
levels of APO4 in an individual. Prostate cancer can also be
g APO4 selective binding agents linked to a therapeutic
polypeptides are also useful for identifying selective
ts, useful in diagnosis/treatment of disease by binding of
e polypeptide/active fragment which is extracellular, or
the cell surface. The binding is preferably performed in
olypeptides/ active fragments are also useful for screening
and antagonists by binding and observing the change in APO4
fective pharmacological agents useful in diagnosis or
disease are also identified using APO4 polypeptides/active
d APO4 signal transducer molecules that specifically interact
lamic domain of APO4 and detecting a change in level of APO4
e method is performed in vivo or in vitro. APO polypeptides
ul as immunogens for preparing antibodies. APO4 is also
agnosis/treatment of developmental or gestational
s. APO8 was transfected to human breast carcinoma cell line
duced apoptosis

AA;

16.2%; Score 46; DB 2; Length 211;
larity 100.0%; Pred. NO. 9.6e-34;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LRLCQVSGLLALRPGSSLRITLPWAHLKAAPFLTYGLFQVH 284
|||||
LRLCQVSGLLALRPGSSLRITLPWAHLKAAPFLTYGLFQVH 211

adard; protein; 225 AA.

(first entry)

tumour necrosis factor related ligand (TRELL).

r necrosis factor related ligand; tnfr; treatment; cancer;
isease; immune system; stimulation; suppression;
ion.

Location/Qualifiers
1. 21
/note= "hydrophobic, transmembrane domain"

97WO-US013945.

96US-0023541P.

96US-0028515P.

97US-0040820P.

EN INC.
GENEVA FACULTY MEDICINE.

PI Chicheportiche Y, Browning JL;
XX WPI; 1998-145619/13.
DR N-PSDB; AAV18539.
XX
PT Tumour necrosis factor related ligand - useful for, e.g. treatir
PT auto-immune disease and immune responses to tissue grafts.
XX
XX Claim 12; Page 48-50; 69pp; English.

XX The sequence is that of mouse tumour necrosis factor related lig
CC (TRELL). TRELL or active fragments can be included with a carrie
CC pharmaceutical compositions to treat cancer, autoimmune diseases
CC immune responses to tissue grafts, or to stimulate or suppress t
CC system. It is useful to screen for TRELL receptors, by labelling
CC detectable label and screening compositions for binding. Agents
CC interfering with TRELL-receptor binding can also be screened for
CC then be administered, optionally with interferon- gamma, to indu
CC death or treat, suppress or alter immune responses (especially i
CC human adenocarcinoma cells) involving a signal pathway between t
CC its receptor. It's coding sequence can be used in gene therapy f
CC related disorders in mammals (especially humans), e.g. tumours,
CC autoimmune and inflammatory diseases or inherited genetic disord
CC introducing into cells, and expressing, therapeutically effectiv
CC of a vector, e.g. a virus comprising a gene encoding TRELL. It m
CC be of use in the preparation of prepare probes for screening
CC natural/synthetic DNAs for TRELL-encoding sequences and for anti
CC therapy
XX
XX Sequence 225 AA;

Query Match 11.3%; Score 32; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 7.6e-21;
Matches 32; Conservative 0; Mismatches 0; Indels 0;

QY 139 RRAIAAHYEVHPRPGDGAQAGVDGTVSGWEE 170
|||||
DB 80 RRAIAAHYEVHPRPGDGAQAGVDGTVSGWEE 111

RESULT 17

AAB07527

ID AAB07527 standard; protein; 225 AA.

XX AAB07527;

XX 20-OCT-2000 (first entry)

DE Amino acid sequence of a soluble recombinant murine TWEAK protei

XX TWEAK protein; immunological disorder; immune response; inflama

KW TWEAK blocking agent; autoimmune disease; organ transplant rejec

KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock

XX Mus sp.

XX WO200042073-A1.

XX 20-JUL-2000.

XX 14-JAN-2000; 2000WO-US001044.

XX 15-JAN-1999; 99US-0116168P.

XX (BIOJ) BIOGEN INC.

XX Rennert P;

XX WPI; 2000-476036/41.

XX Preventing and treating immune responses using modulators, espec:
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful:
PT treating e.g. inflammation and graft versus host disease.

ig 1; 45pp; English.

sequence represents a TWEAK protein. The specification method for preventing or treating an immunological disorder using an immune response in an animal. The method comprises a TWEAK blocking agent. The method may be used for a treating immune disorders associated with inappropriate d/or activity of TWEAK. These disorders include autoimmune te and chronic inflammation, organ transplant rejection, Host disease (GVHD), lymphoid cell malignancies, septic and f shock, loss of immune responsiveness (as seen in human ncy virus (HIV) infections) and failure of the immune umour growth

AA;

11.3%; Score 32; DB 3; Length 225;
arity 100.0%; Pred. No. 7.6e-21;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

AAHYEHPRGQGAQAGVDGTVSGWEE 170

|||||

AAHYEHPRGQGAQAGVDGTVSGWEE 111

lard; protein; 249 AA.

(first entry)

AK.

AK; TNF relatedness and weak ability to induce cell death;
scrosis Factor; TWEAK; fibrosis; cardiac disease;
lung disease; kidney disease; skin disease;
le disease; adipose tissue disease;
nal tract disease; pancreatic disease;
rgan disease; neural disease; cartilage disease;
connective tissue disease; cellular death; hepatotropic;
l; gastrointestinal; osteopathic.

12.

003WO-US011350.

002US-037161P.

INC.

ubowski A, Zheng T, Hahn K;

56/78.

13.

AK-related condition, e.g. liver, gastrointestinal, kidney,
ic, cartilage or neural tissue condition in a subject
nistering to the subject a TWEAK agonist or antagonist.

ID NO 1; 120pp; English.

sequence is murine transmembrane FL-TWEAK (TNF relatedness
ty to induce cell death, where TNF is Tumour Necrosis
is a member of the TNF family. TWEAK agonists or
e useful for treating a TWEAK-related condition, e.g.
liac disease; liver disease; lung disease; kidney disease;

CC skin disease; skeletal muscle disease; adipose tissue disease;
CC gastrointestinal tract disease; pancreatic disease; reproductive
CC disease; neural disease; cartilage disease; bone disease; connect
CC tissue disease; cellular death; and a pathological condition of
CC expressing a TWEAK receptor.

SQ Sequence 249 AA;

Query Match 11.3%; Score 32; DB 7; Length 249;

Best Local Similarity 100.0%; Pred. No. 8.4e-21;
Matches 32; Conservative 0; Mismatches 0; Indels 0; G

Qy 139 RRAIAHYEHPRGQGAQAGVDGTVSGWEE 170

|||||

Db 104 RRAIAHYEHPRGQGAQAGVDGTVSGWEE 135

RESULT 19

AAG01265

ID AAG01265 standard; protein; 58 AA.

AC AAG01265;

XX

DT 06-OCT-2000 (first entry)

XX

DE Human secreted protein, SEQ ID NO: 5346.

XX

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA iso
KW gene therapy; chromosome mapping.

XX

OS Homo sapiens.

XX

PN EP1033401-A2.

XX

PD 06-SEP-2000.

XX

PF 21-FEB-2000; 2000EP-00200610.

XX

PR 26-FEB-1999; 99US-0122487P.

XX

PA (GEST) GENSET.

XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX

DR WPI: 2000-50038-/45.

XX

DR N-PSDB; AAC01272.

XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and fo
PT diagnostic, forensic, gene therapy and chromosome mapping procedu

XX

PS Claim 13; SEQ ID NO 5346; 71pp + Sequence Listing; English.

XX

CC The present sequence is a polypeptide encoded by one of a large n
CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' EST
CC prepared from total human RNAs or polyA+ RNAs derived from 30 dif
CC tissues. EST sequences usually correspond mainly to the 3' untran
CC region (UTR) of the mRNA because they are often obtained from Oli
CC primed cDNA libraries. Such ESTs are not well suited for isolatin
CC sequences derived from the 5' ends of mRNAs and even in those cas
CC longer cDNA sequences have been obtained, the full 5' UTR is rare
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5'
CC are also used in diagnostic, forensic, gene therapy and chromosom
CC mapping procedures. They are used to obtain upstream regulatory s
CC and to design expression and secretion vectors

SQ Sequence 58 AA;

Query Match

Best Local Similarity 3.2%; Score 9; DB 3; Length 58;

Matches 9; Conservative 0; Mismatches 0; Indels 0; G

pecification, but was obtained in electronic format directly
ftp.wipo.int/pub/published_pct_sequences

AA;

3.2%; Score 9; DB 6; Length 365;
arity 100.0%; Pred. No. 17;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

GRGE 50

|||||

GRGE 30

dard; protein; 748 AA.

(first entry)

o acid sequence for GVs-9.

vaccae protein; antigen; T cell activation; cytokine;
1 maturation; infectious disease; immune disorder; cancer;
ystem; mycobacterial infection; allergy; tuberculosis;
oidosis; lung cancer; asthma; skin disorder; psoriasis;
czema; alopecia areata; skin cancer; basal carcinoma;
carcinoma; melanoma.

vaccae.

98WO-NZ000189.

97US-00996624.

97US-00997080.

97US-00997362.

98US-00095855.

98US-00156181.

98US-00205436.

IS RES & DEV CORP LTD.

n J, Visser ES, Skinner MA, Prestidge RL;

163/36.

368.

une response to an antigen.

209-210; 243pp; English.

provides heat-killed Mycobacterium vaccae, or recombinant
teins. The M. vaccae proteins may be employed to activate T
ural killer cells, to stimulate the production of cytokines,
e expression of co-stimulatory molecules on dendritic cells
, and to enhance dendritic cell maturation and function. The
be expressed by standard recombinant methodology.

1 compositions comprising the proteins or nucleic acid
oding the proteins can be used for the treatment,
nd detection of disorders including infectious diseases,
ers and cancer. In particular, the compounds and methods are
tment of diseases of the respiratory system, such as
infections, asthma, allergies, tuberculosis, leprosy,
nd lung cancers, and disorders of the skin such as
opic dermatitis, eczema, allergic contact dermatitis,
ta, and skin cancers such as basal carcinoma, squamous cell
melanoma

XX SQ Sequence 748 AA;

Query Match 3.2%; Score 9; DB 2; Length 748;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

QY 60 ALGLGLALA 68

|||||

Db 282 ALGLGLALA 290

RESULT 23

ABB73512

ID ABB73512 standard; protein; 749 AA.

XX ABB73512;

AC ABB73512;

XX 08-APR-2002 (first entry)

XX M vaccae GVs-9 protein SEQ ID NO: 154.

XX Skin disorder; psoriasis; atopic dermatitis; allergic contact der;
XX alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cyt
XX antiporiatic; dermatological; antiinflammatory; antiallergic;
XX Th2 immune response; immunomodulatory.

OS Mycobacterium vaccae.

XX US6328978-B1.

XX 11-DEC-2001.

XX 02-JUN-1999; 99US-00324542.

XX 23-DEC-1997; 97US-00997080.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Tan PLJ, Prestidge R;

XX WPI; 2002-138361/18.

XX N-PSDB; ABL36274.

XX Inhibiting skin inflammation associated with skin disorder e.g.
XX psoriasis, by administering composition comprising delipidated ar
XX deglycolipidated Mycobacterium vaccae cells or Mycobacterium vacc
XX culture filtrate.

XX Example 6; Col 161-164; 116pp; English.

XX The present invention relates to a method of inhibiting skin infl
XX associated with a skin disorder selected from psoriasis, atopic
XX dermatitis and allergic contact dermatitis, which involves admini
XX a composition containing delipidated and deglycolipidated Mycobac
XX vaccae cells or M. vaccae culture filtrate. The skin disorder to
XX treated may also include alopecia areata, and skin cancers such a
XX cell carcinoma, squamous cell carcinoma and melanoma. The composi
XX acts by inhibiting the Th2 immune response. The present sequence
XX protein described in the exemplification of the invention

XX SQ Sequence 749 AA;

Query Match 3.2%; Score 9; DB 5; Length 749;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; G

QY 60 ALGLGLALA 68

|||||

Db 282 ALGLGLALA 290

RESULT 24

standard; protein; 54 AA.

(first entry)

erium acnes immunogenic protein #12759.

me; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
ophthalmitis; bone; joint; central nervous system; ELISA;
lesion; acne vulgaris; enzyme linked immunosorbent assay;
al; osteopathic; neuroprotectant.

erium acnes.

A2.

2001WO-US012865.

2000US-0199047P.

2000US-0208841P.

2000US-0216747P.

XA CORP.

Persing DH, Mitcham JL, Wang SS, Bhatia A;
e J, Zhang Y, Jen S, Carter D;

6774/71.

9552.

erium acnes polypeptides and nucleic acids useful for
against and diagnosing infections, especially useful for
e vulgaris.

EQ ID NO 13058; 1069pp; English.

U39105-AA068017 represent Propionibacterium acnes immunogenic
t. The proteins and their associated DNA sequences are used in
t, prevention and diagnosis of medical conditions caused by
e disorders include SAPHO syndrome (synovitis, acne,
hypertosis and osteomyelitis), uveitis and endophthalmitis.
also involved in infections of bone, joints and the central
em, however it is particularly involved in the inflammatory
ciated with acne vulgaris. A method for detecting the
absence of P. acnes in a patient comprises contacting a
binding agent that binds to the proteins of the invention
ing the amount of bound protein in the sample. The
may be used as antigens in the production of antibodies
P. acnes proteins. These antibodies can be used to
expression and activity of P. acnes polypeptides and
eat P. acnes infections. The antibodies may also be used as
gents for determining P. acnes presence, for example, by
d immunosorbent assay (ELISA). Note: The sequence data for
did not form part of the printed specification, but was
electronic format directly from WIPO at
/pub/published_pct_sequences

AA;

larity 2.8%; Score 8; DB 4; Length 54;
100.0%; Pred. No. 23;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PLPRS 18

|||||

PLPRS 28

ABM48382

ID ABM48382 standard; protein; 54 AA.

XX AC ABM48382;

XX DT 20-OCT-2003 (first entry)

XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #13058
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW Immunostimulant; immune response; vaccine.

XX OS Propionibacterium acnes.

XX PN WO2003033515-A1.

XX PD 24-APR-2003.

XX PF 11-OCT-2002; 2002WO-US032727.

XX PR 15-OCT-2001; 2001US-00978825.

XX PA (CORI-) CORIXA CORP.

XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JI
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Cart
PI Barth B, Vallieve-Douglas J;

XX DR WPI; 2003-381789/36.

DR N-PSDB; ACF64481.

XX PT New Propionibacterium acnes polypeptides and polynucleotides enc
PT polypeptide, useful for diagnosing, preventing or treating acne
PT or for stimulating an immune response specific for a P. acnes pr

PS Example 1; SEQ ID NO 13058; 1481pp; English.

XX CC The invention relates to an isolated polynucleotide (ACF64435-AC
CC encoding a Propionibacterium acnes protein. The invention also r
CC polypeptides encoded by the polynucleotides (ABM35624-ABM4536)
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells compr
CC polynucleotide of the invention; antibodies against polypeptides
CC invention; fusion proteins comprising a polypeptide of the inven
CC method for stimulating an immune response specific for a P. acne
CC polypeptide and an isolated T cell population comprising T cells
CC via this method; a vaccine composition (comprising P. acnes poly
CC polynucleotides, antibodies, fusion proteins, T cell populations
CC antigen-presenting cells that express the polypeptide); a method
CC for detecting or determining the presence or absence of P. acnes
CC patient; and a method for inhibiting the development of P. acnes
CC patient. The P. acnes polypeptides, polynucleotides, antibodies,
CC proteins, T cell populations or antigen-presenting cells that ex
CC polypeptides are useful for diagnosing, preventing or treating a
CC vulgaris, or for stimulating an immune response specific for a P
CC protein. The polynucleotides can also be used as probes or prime
CC nucleic acid hybridisation. The vaccine composition is useful fo
CC stimulation of an immune response against P. acnes, or for treat
CC and the kit is useful for performing a diagnostic assay. The pre
CC sequence represents a polypeptide predicted to be encoded by an
CC reading frame) contained within the P. acnes polynucleotides of
CC invention. Note: The sequence data for this patent did not form
CC the printed specification, but was obtained in electronic format
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 54 AA;

Query Match 2.8%; Score 8; DB 6; Length 54;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 11 RRLPLPRS 18

|||||

LPRS 28

dard; protein; 55 AA.

(first entry)

encoded by probe for measuring cervical gene expression.
microarray; gene expression; cervical epithelial cell;
er.

2.

2001WO-US000670.

2000US-0180312P.
2000US-0207456P.
2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

ULAR DYNAMICS INC.

zel DK, Chen W, Rank DR;

901/53.

derived single exon nucleic acid probes useful for analyzing
on in human cervical epithelial cells.

ID NO 25684; 487pp; English.

vention relates to human single exon nucleic acid probes
110068-AA128459). The present sequence is a peptide encoded
robe. The SNPs are derived from human HeLa cells. The SNPs
o produce a single exon microarray, which can be used for
an gene expression in a sample derived from human cervical
lles. By measuring gene expression, the probes are therefore
ing and/or staging of diseases of the cervix, notably
er. Note: The sequence data for this patent did not form
rinted specification, but was obtained in electronic format
WIPO at ftp.wipo.int/pub/published_pct_sequences

A;

arity 2.8%; Score 8; DB 4; Length 55;
onservative 100.0%; Pred.No. 23;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

LALA 68
|
|
|
LALA 19

dard; peptide; 55 AA.

(first entry)

DE Peptide #10143 encoded by human foetal liver single exon probe.
XX Human; foetal liver; gene expression; single exon nucleic acid p;
XX Homo sapiens.
OS WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
PT gene expression in human fetal liver.
XX Claim 27; SEQ ID NO 35272; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for mea
CC single gene expression in a sample derived from human foetal liver;
CC single exon nucleic acid probes may be used for predicting, measu
CC displaying gene expression in samples derived from human fetal li
CC present sequence is a peptide encoded by a single exon nucleic ac
CC of the invention. Note: The sequence data for this patent did not
CC part of the printed specification, but was obtained in electronic
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 55 AA;
SQ

Query Match 2.8%; Score 8; DB 4; Length 55;
Best Local Similarity 100.0%; Pred.No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 61 LGUGLALA 68
Db 12 LGUGLALA 19
|
|
|
|

RESULT 28
AAM36451
ID AAM36451 standard; protein; 55 AA.
XX

AC AAM36451;

DT 17-OCT-2001 (first entry)

DE Peptide #10488 encoded by probe for measuring placental gene expr
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX Homo sapiens.

OS WO200157272-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000663.
XX 04-FEB-2000; 2000US-0180312P.
PR

2000US-0207456P.
 2000US-00608408.
 2000US-00632366.
 2000US-0234687P.
 2000US-0236359P.
 2000GB-00024263.
 MOLECULAR DYNAMICS INC.
 nzel DK, Chen W, Rank DR;
 8897/53.
 -derived single exon nucleic acid probes useful for analyzing
 ion in human placenta.
 Q ID NO 36720; 654pp; English.
 invention relates to single exon nucleic acid probes (SENP:
 -AA157546). The present sequence is a peptide encoded by one
 measuring and displaying gene expression in samples derived
 placenta. The probes are useful for antenatal diagnosis of
 c disorders
 AA;
 larity 2.8%; Score 8; DB 4; Length 55;
 Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
 GLALA 68
 |||||
 GLALA 19
 ndard; protein; 55 AA.
 (first entry)
 2 encoded by probe for measuring heart cell gene expression.
 expression; heart; microarray; vascular system;
 ar disease; hypertension; cardiac arrhythmia;
 eart disease.
 A2.
 2001WO-US0000666.
 2000US-0180312P.
 2000US-0207456P.
 2000US-00608408.
 2000US-00632366.
 2000US-0234687P.
 2000US-0236359P.
 2000GB-00024263.
 CULAR DYNAMICS INC.
 nzel DK, Chen W, Rank DR;
 8899/53.
 nucleic acid probes for analyzing gene expression in human

PT hearts.
 XX Claim 15; SEQ ID NO 27753; 530pp; English.
 XX The present invention relates to single exon nucleic acid probes
 CC measuring human gene expression in a sample derived from human
 CC ABA21535-ABA41305). The present sequence is a protein encoded by
 CC probe. The probes may be used for predicting, measuring and dis
 CC gene expression in samples derived from the human heart via mic
 CC By measuring gene expression, the probes are useful for predict
 CC diagnosing, grading, staging, monitoring and prognosing disease
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly f
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 55 AA;
 Query Match 2.8%; Score 8; DB 4; Length 55;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservativity 0; Mismatches 0; Indels 0;
 Qy 61 LGGLGALA 68
 Db 12 LGGLGALA 19
 RESULT 30
 AAM76342
 ID AAM76342 standard; protein; 55 AA.
 XX
 AC AAM76342;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36
 KW Human; bone marrow expressed exon; gene expression analysis; pr
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000668.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 gene expression in human bone marrow.
 XX
 PS Example 4; SEQ ID NO 36648; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic a
 CC probes which are derived from genomic sequences expressed in the
 CC bone marrow. They can be used to measure gene expression in bone
 CC samples, which may enable the improved diagnosis and treatment o
 CC such as lymphoma, leukaemia and myeloma. The present sequence is

by one of the probes of the invention

2.8%; Score 8; DB 4; Length 55;
 100.0%; Pred. No. 23;
 0; Mismatches 0; Indels 0; Gaps 0;
 ALA 68
 ||||
 ALA 19

ard; protein; 55 AA.

(first entry)

pressed single exon probe encoded protein SEQ ID NO: 35633.
 expressed exon; gene expression analysis; probe; microarray;
 disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

2.

2001WO-US000667.

2000US-0180312P.
 2000US-0207456P.
 2000US-00608408.
 2000US-00632366.
 2000US-0234687P.
 2000US-0236359P.
 2000GB-00024263.

ULAR DYNAMICS INC.

zel DK, Chen W, Rank DR;

446/52.

ucleic acid probes for analyzing gene expression in human

Q ID NO 35633; 650pp + Sequence Listing; English.

vention provides a number of single exon nucleic acid
 are derived from genomic sequences expressed in the human
 an be used to measure gene expression in brain cell samples,
 ble the diagnosis and improved treatment of nervous system
 as Alzheimer's disease, multiple sclerosis, schizophrenia, f
 cancers. The present sequence is a protein encoded by one of
 the invention

2.8%; Score 8; DB 4; Length 55;
 100.0%; Pred. No. 23;
 0; Mismatches 0; Indels 0; Gaps 0;
 ALA 68
 ||||
 ALA 19

ABG58050
 ID ABG58050 standard; peptide; 55 AA.

AC ABG58050;
 XX 25-FEB-2003 (first entry)
 DE Human liver peptide, SEQ ID No 36698.
 XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 XX hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000664.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for a
 gene expression in human adult liver.

Claim 27; SEQ ID NO 36698; 659pp; English.

The invention relates to a single exon nucleic acid probe (SENP)
 measuring human gene expression in a sample derived from human ac
 liver, comprising one of 13109 defined nucleotide sequences giver
 specification (or complements/ fragments). The probe hybridises a
 stringency to a nucleic acid molecule expressed in the human adul
 (I) may be used for predicting, measuring and displaying gene exp
 in samples derived from human adult liver. The genes identified
 involved in genetic liver diseases such as cirrhosis,
 hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia
 associated with coronary heart disease. ABG47348-ABG59930 repres
 liver single exon encoded peptides of the invention. Note: The se
 information for this patent does not appear in the printed speci
 but was obtained in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

XX Sequence 55 AA;

Query Match 2.8%; Score 8; DB 4; Length 55;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

OY 61 LGLGLALA 68
 |||||
 Db 12 LGLGLALA 19

RESULT 33

ABG45635
 ID ABG45635 standard; peptide; 55 AA.

XX AC ABG45635;

XX DT 19-AUG-2002 (first entry)

XX

e encoded by genome-derived single exon probe SEQ ID 35300.
 e exon probe; asthma; lung cancer; COPD; ILD;
 ractive pulmonary disease; interstitial lung disease;
 opathic pulmonary fibrosis; neurofibromatosis;
 erosis; Gaucher's disease; Niemann-Pick disease;
 dlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 sticyctosis; lymphangioleiomyomatosis; Karagener syndrome;
 veolar proteinosis; fibrocystic pulmonary dysplasia;
 ary dyskinesia; pulmonary hypertension;
 rane disease.

A2.

2001WO-US000665.

2000US-0180312P.

2000US-0207456P.

2000US-00608408.

2000US-00632366.

2000US-0234687P.

2000US-0236359P.

2000GB-00024263.

TULAR DYNAMICS INC.

Izel DK, Chen W, Rank DR;

1183/15.

Irreversible set of single exon nucleic acid probes, used to
 expression in human lung samples.

2 ID NO 35300; 634pp; English.

1 relates to a spatially-addressable set of single exon
 probes for measuring gene expression in a sample deriv
 ing comprising single exon nucleic acid probes having one of
 ; acid sequences mentioned in the specification, or their
 or the 12387 open reading frames derived from the 12614
 included are a microarray comprising the novel set of probes
 set of probes which hybridise at high stringency to a nucleic
 id in the human lung; measuring gene expression in a sample
 human lung, comprising (a) contacting the array with a
 ; detectably labeled nucleic acids derived from human lung
 measuring the label detectably bound to each probe of the
 .fying exons in a eukaryotic genome, comprising (a)
 ly predicting at least one exon from genomic sequences of
 ; and (b) detecting specific hybridisation of detectably
 ic acids from eukaryotic lung mRNA, to a single exon probe,
 ment identical to the predicted exon, the probe is included
 mentioned microarray; assigning exons to a single gene,
) identifying exons from genomic sequence by the method
 measuring the expression of each of the exons in several
 or cell types using hybridisation to a single exon
 having a probe with the exon, where a common pattern of
 ; the exons in the tissues and/or cell types indicates that
 uld be assigned to a single gene; a peptide comprising one
 ences, mentioned in the specification, or encoded by the
 eading frames (ORF). The probes are used for gene expression
 l for identifying exons in a gene, particularly using human
 mRNA and for the study of lung diseases such as asthma, lung
 ic obstructive pulmonary disease (COPD), interstitial lung
 , familial idiopathic pulmonary fibrosis, neurofibromatosis,
 xosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 me, sarcoidosis, pulmonary haemosiderosis, pulmonary
 , lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 idrome, fibrocystic pulmonary dysplasia, primary ciliary
 pulmonary hypertension and hyaline membrane disease. The

CC present sequence is a peptide/protein encoded by a single exon I
 CC the invention. Note: The sequence data for this patent did not i
 CC of the printed specification, but was obtained in electronic for
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 55 AA;

Query Match 2.8%; Score 8; DB 5; Length 55;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 61 LGLGLALA 68
 Db 12 LGLGLALA 19
 |||||
 |||||

RESULT 34

AA21621
 ID AA21621 standard; protein; 65 AA.

XX

AC

DT 12-OCT-2001 (first entry)

XX

DE Peptide #8055 encoded by probe for measuring cervical gene expre
 XX Probe; human; microarray; gene expression; cervical epithelial c
 KW cervical cancer.
 XX Homo sapiens.

OS

XX

PN

XX

PD

XX

PF 30-JAN-2001; 2001WO-US000670.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

XX WPI; 2001-488901/53.

DR

PT Human genome-derived single exon nucleic acid probes useful for
 PT gene expression in human cervical epithelial cells.

XX

PS Claim 27; SEQ ID NO 26447; 487pp; English.

XX

CC The present invention relates to human single exon nucleic acid I
 CC (SENP; see AA110068-AA128459). The present sequence is a peptide
 CC by one such probe. The SENPs are derived from human HeLa cells.
 CC can be used to produce a single exon microarray, which can be use
 CC measuring human gene expression in a sample derived from human ce
 CC epithelial cells. By measuring gene expression, the probes are ti
 CC useful in grading and/or staging of diseases of the cervix, notat
 CC cervical cancer. Note: The sequence data for this patent did not
 CC part of the printed specification, but was obtained in electronic
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 65 AA;

Query Match

Best Local Similarity 2.8%; Score 8; DB 4; Length 65;

Matches 8; Conservative 0; Mismatches 0; Indels 0; G

2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

ULAR DYNAMICS INC.

zel DK, Chen W, Rank DR;

3899/53.

nucleic acid probes for analyzing gene expression in human

Q ID NO 28660; 530pp; English.

invention relates to single exon nucleic acid probes for
nan gene expression in a sample derived from human heart (see
11305). The present sequence is a protein encoded by one such
cobs may be used for predicting, measuring and displaying
ion in samples derived from the human heart via microarrays.
gene expression, the probes are useful for predicting, the
grading, staging, monitoring and prognosing diseases of the
and vascular system e.g. cardiovascular disease,
cardiac arrhythmias and congenital heart disease. Note: The
a for this patent did not form part of the printed
n, but was obtained in electronic format directly from WIPO
nt./pub/published_pct_sequences

LA;

2.8%; Score 8; DB 4; Length 65;
arity 100.0%; Pred. No. 27;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALA 68
|||
ALA 29

ard; protein; 65 AA.

(first entry)

arrow expressed probe encoded protein SEQ ID NO: 38012.

arrow expressed exon; gene expression analysis; probe;
ancer; leukaemia; lymphoma; myeloma.

2.

2001WO-US000668.

2000US-0180312P.
2000US-0207456P.
2000US-00608408.
2000US-00632366.
2000US-0234687P.
2000US-0236359P.
2000GB-00024263.

ULAR DYNAMICS INC.

zel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488900/53.

XX

Human genome-derived single exon nucleic acid probes useful for
gene expression in human bone marrow.

XX

PS Example 4; SEQ ID NO 38012; 658pp + Sequence Listing; English.

XX

The present invention provides a number of single exon nucleic a
probes which are derived from genomic sequences expressed in the
bone marrow. They can be used to measure gene expression in bone
samples, which may enable the improved diagnosis and treatment o
such as lymphoma, leukaemia and myeloma. The present sequence is
protein encoded by one of the probes of the invention

XX

SQ Sequence 65 AA;

Query Match 2.8%; Score 8; DB 4; Length 65;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 61 LGLGLALA 68

Db 22 LGLGLALA 29

RESULT 39

AAM64984

XX

ID AAM64984 standard; protein; 65 AA.

XX

AC AAM64984;

XX

DT 05-NOV-2001 (first entry)

XX

DE Human brain expressed single exon probe encoded protein SEQ ID N

XX

KW Human; brain expressed exon; gene expression analysis; probe; mi

KW

Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy

XX

OS Homo sapiens.

XX

PN WO200157275-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000667.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR

26-MAY-2000; 2000US-0207456P.

PR

30-JUN-2000; 2000US-00608408.

PR

03-AUG-2000; 2000US-00632366.

PR

21-SEP-2000; 2000US-0234687P.

PR

27-SEP-2000; 2000US-0236359P.

PR

04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

WPI; 2001-483446/52.

XX

Single exon nucleic acid probes for analyzing gene expression in

PT

brains.

XX

Example 4; SEQ ID NO 37089; 650pp + Sequence Listing; English.

XX

The present invention provides a number of single exon nucleic ac
probes which are derived from genomic sequences expressed in the
brain. They can be used to measure gene expression in brain cell
which may enable the diagnosis and improved treatment of nervous
diseases such as Alzheimer's disease, multiple sclerosis, schizof
epilepsy and cancers. The present sequence is a protein encoded t
the probes of the invention

2.8%; Score 8; DB 4; Length 65;
 100.0%; Pred. No. 27;
 0; Mismatches 0; Indels 0; Gaps 0;

ALA 68
 ||||
 ALA 29

lard; peptide; 65 AA.

first entry)

ptide, SEQ ID No 38009.

cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 olaemia; coronary heart disease.

001WO-US000664.

000US-0180312P.
 000US-0207456P.
 000US-00608408.
 000US-00632366.
 000US-0234687P.
 000US-0236359P.
 000GB-00024263.

LAR DYNAMICS INC.

el DK, Chen W, Rank DR;
 98/53.

erived single exon nucleic acid probes useful for analyzing
 n in human adult liver.

ID NO 38009; 659pp; English.

relates to a single exon nucleic acid probe (SENP) (I) for
 n gene expression in a sample derived from human adult
 ing one of 13109 defined nucleotide sequences given in the
 (or complements/ fragments). The probe hybridises at high
 a nucleic acid molecule expressed in the human adult liver.
 i for predicting, measuring and displaying gene expression
 ived from human adult liver. The genes identified may be
 netic liver diseases such as cirrhosis,
 naemia, hyperlipidaemia and hypercholesterolaemia which is
 h coronary heart disease. ABG47348-ABG59930 represent human
 xon encoded peptides of the invention. Note: The sequence
 r this patent does not appear in the printed specification
 ed in electronic format directly from WIPO at
 ub/published_pct_sequences

2.8%; Score 8; DB 4; Length 65;
 100.0%; Pred. No. 27;
 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGLGLALA 68
 |||||
 Db 22 LGLGLALA 29

RESULT 41
 ABG46737
 ID ABG46737 standard; peptide; 65 AA.
 XX
 AC ABG46737;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE
 XX
 KW Human peptide encoded by genome-derived single exon probe SEQ ID
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndr
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW Primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used
 PT measure gene expression in human lung samples.
 XX
 PS Claim 27; SEQ ID NO 36402; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exc
 CC nucleic acid probes for measuring gene expression in a sample deri
 CC from human lung comprising single exon nucleic acid probes having
 CC 12614 nucleic acid sequences mentioned in the specification, or th
 CC complements or the 12387 open reading frames derived from the 1261
 CC probes. Also included are a microarray comprising the novel set of
 CC ; the novel set of probes which hybridise at high stringency to a
 CC acid expressed in the human lung; measuring gene expression in a s
 CC derived from human lung, comprising (a) contacting the array with
 CC collection of detectably labeled nucleic acids derived from human
 CC mRNA, and (b) measuring the label detectably bound to each probe o
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequence
 CC the eukaryote; and (b) detecting specific hybridisation of detecta
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon p
 CC having a fragment identical to the predicted exon, the probe is in
 CC in the above mentioned microarray; assigning exons to a single gen
 CC comprising (a) identifying exons from genomic sequence by the meth
 CC above and (b) measuring the expression of each of the exons in sev
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern o

the exons in the tissues and/or cell types indicates that could be assigned to a single gene; a peptide comprising one or more amino acids, mentioned in the specification, or encoded by the reading frames (ORF). The probes are used for gene expression analysis for identifying exons in a gene, particularly using human cDNA and for the study of lung diseases such as asthma, lung cancer, obstructive pulmonary disease (COPD), interstitial lung disease, familial idiopathic pulmonary fibrosis, neurofibromatosis, Crohn's disease, Gaucher's disease, Niemann-Pick disease, Hermansky-Jaeger syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary lymphangioleiomyomatosis, pulmonary alveolar proteinosis, idiopathic pulmonary fibrosis, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The invention is a peptide/protein encoded by a single exon probe of the invention. The sequence data for this patent did not form part of the specification, but was obtained in electronic format from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 2.8%; Score 8; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 0; Mismatches 0; Indels 0; Gaps 0;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
HLALA 68
|||||
HLALA 29

Standard; protein; 69 AA.

(first entry)

Diagnostic protein #3714.

Gene mapping; gene mapping; gene therapy; forensic;
Medical imaging; diagnostic; genetic disorder.

2.

2001WO-US008631.

2000US-00540217.

2000US-00649167.

INC.

Liu C, Tang YT;

362/73.

910.

polynucleotide and encoded polypeptides, useful in
diagnostics, gene mapping, identification of mutations
or genetic disorders or other traits and to assess

ID NO 34082; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene ma
in recombinant production of (II). The polynucleotides are al
in diagnostics as expressed sequence tags for identifying express
genes (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II)
useful in gene therapy techniques to restore normal

activity of (II) or to treat disease states involving (II). (II)
useful for generating antibodies against it, detecting or quantiti
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical
of sites expressing (II). (I) and (II) are useful for treating d
involving aberrant protein expression or biological activity. Th
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutation
responsible for genetic disorders or other traits to assess biod
and to produce other types of data and products dependent on DNA
amino acid sequences. ABG00010-ABG030377 represent novel human di
amino acid sequences of the invention. Note: The sequence data f
patent did not appear in the printed specification, but was obta
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 69 AA;

Query Match 2.8%; Score 8; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

OY 61 LGLGLALA 68
|||||
DB 16 LGLGLALA 23

RESULT 43

ABG03663
ID ABG03663 standard; protein; 71 AA.

AC ABG03663;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #3654.

Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS67850.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.

Claim 20; SEQ ID NO 34022; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypept
sequences. (I) is useful as hybridisation probes, polymerase chai
reaction (PCR) primers, oligomers, and for chromosome and gene ma
in recombinant production of (II). The polynucleotides are al
in diagnostics as expressed sequence tags for identifying express
genes (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II)
useful for generating antibodies against it, detecting or quantit

n tissue, as molecular weight markers and as a food II) and its binding partners are useful in medical imaging assays (II). (I) and (II) are useful for treating disorders and protein expression or biological activity. The and polynucleotide sequences have applications in forensics, gene mapping, identification of mutations or genetic disorders or other traits to assess biodiversity a other types of data and products dependent on DNA and sequences. ABG00010-ABG30377 represent novel human diagnostic features of the invention. Note: The sequence data for this t appear in the printed specification, but was obtained in rmat directly from WIPO at pub/published_pct_sequences

A;
2.8%; Score 8; DB 4; Length 71;
arity 100.0%; Pred. No. 29; 0; Indels 0; Gaps 0;
nservative 0; Mismatches 0; Mismatches 0; Indels 0; Gaps 0;
ALA 68
|||
ALA 23

lard; protein; 84 AA.

(first entry)

ium acnes immunogenic protein #22178.

; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
hthalmitis; bone; joint; central nervous system; ELISA;
esion; acne vulgaris; enzyme linked immunosorbent assay;
; osteopathic; neuroprotectant.

ium acnes.

001WO-US012865.

000US-0199047P.

000US-0208841P.

000US-0216747P.

CORP.

ersing DH, Mitcham JL, Wang SS, Bhatia A;
J, Zhang Y, Jen S, Carter D;

74/71.

15.

ium acnes polypeptides and nucleic acids useful for
ainst and diagnosing infections, especially useful for
vulgaris.

ID NO 22477; 1069pp; English.

9105-AAU68017 represent Propionibacterium acnes immunogenic
The proteins and their associated DNA sequences are used in
prevention and diagnosis of medical conditions caused by
disorders include SAPHO syndrome (synovitis, acne,
petrosis and osteomyelitis), uveitis and endophthalmitis.
so involved in infections of bone, joints and the central
so, however it is particularly involved in the inflammatory

lesions associated with acne vulgaris. A method for detecting the
presence or absence of P. acnes in a patient comprises contacting
sample with a binding agent that binds to the proteins of the in
and determining the amount of bound protein in the sample. The
polypeptides may be used as antigens in the production of antibod
specific for P. acnes proteins. These antibodies can be used to
downregulate expression and activity of P. acnes polypeptides and
therefore treat P. acnes infections. The antibodies may also be u
diagnostic agents for determining P. acnes presence, for example,
enzyme linked immunosorbent assay (ELISA). Note: The sequence dat
this patent did not form part of the printed specification, but w
obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX Sequence 84 AA;

Query Match 2.8%; Score 8; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 259 LRITLPM 266

Db 34 LRITLPM 41

RESULT 45

ABM57801

ID ABM57801 standard; protein; 84 AA.

XX AC ABM57801;

XX DT 20-OCT-2003 (first entry)

XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #22477.

XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.

XX OS Propionibacterium acnes.

XX PN WO2003033515-A1.

XX PD 24-APR-2003.

XX PF 11-OCT-2002; 2002WO-US032727.

XX PR 15-OCT-2001; 2001US-00978825.

XX PA (CORI-) CORIXA CORP.

XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter
PI Barth B, Vallieve-Douglas J;

XX WPI; 2003-381789/36.

XX DR N-PSDB; ACF64544.

XX PT New Propionibacterium acnes polypeptides and polynucleotides enco
PT polypeptide, useful for diagnosing, preventing or treating acne v
PT or for stimulating an immune response specific for a P. acnes prot
XX Example 1; SEQ ID NO 22477; 1481pp; English.

XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF6
CC encoding a Propionibacterium acnes protein. The invention also rel
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) an
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells compris
CC polynucleotide of the invention; antibodies against polypeptides c
CC invention; fusion proteins comprising a polypeptide of the inventi
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells p
CC via this method; a vaccine composition (comprising P. acnes polype

ies, antibodies, fusion proteins, T cell populations, or
 enting cells that express the polypeptide); a method and kit
 g or determining the presence or absence of P. acnes in a
 method for inhibiting the development of P. acnes in a
 P. acnes polypeptides, polynucleotides, antibodies, fusion
 cell populations or antigen-presenting cells that express the
 are useful for diagnosing, preventing or treating acne
 for stimulating an immune response specific for a P. acnes
 polynucleotides can also be used as probes or primers for
 hybridisation. The vaccine composition is useful for the
 of an immune response against P. acnes, or for treating acne,
 is useful for performing a diagnostic assay. The present
 resents a polypeptide predicted to be encoded by an ORF (open
 e) contained within the P. acnes polynucleotides of the
 te: The sequence data for this patent did not form part of
 specification, but was obtained in electronic format directly
 ftp.wipo.int/pub/published_pct_sequences

AA;

2.8%; Score 8; DB 6; Length 84;
 larity 100.0%; Pred. No. 34;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RTLPW 266

|||||

RTLPW 41

ard; protein; 110 AA.

(first entry)

diagnostic protein #20250.

some mapping; gene mapping; gene therapy; forensic;
 nt; medical imaging; diagnostic; genetic disorder.

12.

2001WO-US008631.

2000US-00540217.

2000US-00649167.

} INC.

Liu C, Tang YT;

1362/73.

.446.

polynucleotide and encoded polypeptides, useful in
 forensics, gene mapping, identification of mutations
 for genetic disorders or other traits and to assess

} ID NO 50618; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
) is useful as hybridisation probes, polymerase chain
) primers, oligomers, and for chromosome and gene mapping,
 nant production of (II). The polynucleotides are also used
 as expressed sequence tags for identifying expressed
 useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II)
 CC useful for generating antibodies against it, detecting or quanti
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical
 CC of sites expressing (II). (I) and (II) are useful for treating d
 CC involving aberrant protein expression or biological activity. Th
 CC polypeptide and polynucleotide sequences have application in
 CC diagnostics, forensics, gene mapping, identification of mutation
 CC responsible for genetic disorders or other traits to assess blood
 CC and to produce other types of data and products dependent on DNA
 CC amino acid sequences. ABO00010-ABG30377 represent novel human di
 CC amino acid sequences of the invention. Note: The sequence data f
 CC patent did not appear in the printed specification, but was obta
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 110 AA;

Query Match 2.8%; Score 8; DB 4; Length 110;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 61 LGLGLALA 68

Db 44 LGLGLALA 51

RESULT 47

AAO08094

ID AAO08094 standard; protein; 117 AA.

XX AAO08094;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 21986.

XX Human; cytokine; cell proliferation; cell differentiation; gene
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesi
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI88025.

XX Isolated nucleic acids and polypeptides, useful for preventing d:
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 21986; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity el
 CC cytokine, cell proliferation or cell differentiation or which may
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vac
 CC peptide therapy. The polypeptides have various cytokine-like acti
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activit

in activity and may be useful in the diagnosis and/or cancer, leukaemia, nervous system disorders, arthritis and Note: The sequence data for this patent did not form part of the specification, but was obtained in electronic format from WIPO at ftp.wipo.int/pub/published_pct_sequences

AA;

2.8%; Score 8; DB 4; Length 117;
arity 100.0%; Pred.No. 47;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

ALA 68
||||
ALA 61

ard; protein; 184 AA.

(first entry)

lanogaster polypeptide SEQ ID NO 29928.

velopmental biology; cell signalling; insecticide;

lanogaster.

).

2001WO-US009231.

2000US-0191637P.

2000US-00614150.

ip NY.

Iams M, Li PWD, Myers EW;

160/75.

115.

nucleic acid detection reagent for detecting 1000 or more nucleophila and for elucidating cell signalling and cell-cell

Q ID NO 29928; 21pp + Sequence Listing; English.

relates to an isolated nucleic acid detection reagent acting 1000 or more genes from Drosophila. The invention is a method for elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of therapeutic and pharmaceutical drugs. The invention relates to nucleic acid sequences (AB116176-ABL30511), expressed DNA sequences (AB116175) and the encoded proteins (AB57737-AB57738). The sequence data for this patent did not form part of the specification, but was obtained in electronic format directly from ftp.wipo.int/pub/published_pct_sequences

AA;

2.8%; Score 8; DB 4; Length 184;
arity 100.0%; Pred.No. 73;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

WSL 79
|||

Db 16 LLLAWSL 23

RESULT 49

ID ABP28041 standard; protein; 190 AA.

XX ABP28041;

XX AC ABP28041;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 5258.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus aga

XX KW Group A streptococcus; Streptococcus pyogenes; antibacterial;

XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB004789.

XX PR 27-OCT-2000; 2000GB-00026333.

XX PR 24-NOV-2000; 2000GB-00028727.

XX PR 07-MAR-2001; 2001GB-00005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Masighani V, Margarit Y RosI, Grandi G, Fraser C;

XX PI Tettelin H;

XX DR WPI; 2002-352536/38.

XX DR N-PSDB; ABN68672.

XX PT New Streptococcus protein for the treatment or prevention of infe

XX PT disease caused by Streptococcus bacteria, such as meningitis, and

XX PT detecting a compound that binds to the protein.

XX PS Claim 1; Page 3689; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group

XX CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptoco

XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1),

XX CC the specification. The proteins have antibacterial and antinflam

XX CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and

XX CC antibodies that bind (I) are used in the manufacture of medicamen

XX CC the treatment or prevention of infection or disease caused by

XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogen

XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a

XX CC biological sample. (I) is used to determine whether a compound bin

XX CC (I). A composition comprising (I) or a nucleic acid encoding (I),

XX CC used as a vaccine or diagnostic composition. The disease caused by

XX CC Streptococcus that is prevented or treated may be meningitis. Nuc

XX CC acid encoding (I) may be used to recombinantly produce (I) and ma

XX CC used in gene therapy. Antibodies to (I) are used for affinity

XX CC chromatography, immunoassays, and distinguishing/identifying

XX CC Streptococcus proteins

XX SQ Sequence 190 AA;

Query Match 2.8%; Score 8; DB 5; Length 190;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 258 SLRIRLTP 265

Db 50 SLRIRLTP 57

ward; protein; 198 AA.

(first entry)

oded protein SEQ ID NO: 1209.

y; pig; cow; fruit fly; Yeast; hamster; macaque; horse;
y; dog; sea urchin; expressed sequence tag; EST;
forensic test; gene mapping; genetic disorder; biodiversity;
nutrition.

2.

2001WO-US002687.

2000US-00491404.

2000US-00617746.

2000US-00631451.

2000US-00663870.

INC.

1 C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
iac RA, Zhang J, Werhman T;
164/51.
1343.

Peptide for treatment of diseases, diagnostics, raising
id research use.

873; 1275pp; English.

vention provides the protein and coding sequences of novel
a variety of organisms, including human, dog, cat, horse,
ater, monkey, macaque, yeast, bacteria, fruit fly, sea
mato. These were derived from expressed sequence tags (ESTs)
nism of interest. They can be used in diagnostics,
ne mapping, identification of mutations, to assess
and for nutritional purposes. The present sequence is a
e invention

AA;

2.8%; Score 8; DB 4; Length 198;
arity 100.0%; Pred. No. 78;
onservative 0; Mismatches 0; Indels 0; Gaps 0;

PLAL 61
|||||
PLAL 9

ward; protein; 222 AA.

(first entry)

antigen HPAMG11, SEQ ID NO:2806.

an antigen; ovary; ovarian; breast; cancer; tumour;
ir; breast cancer; tumour; reproductive system disorder;

infertility; pregnancy disorder; anovulation; polycystic ovary s
PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection
inflammatory condition; immune disorder; blood disorder;
cardiovascular disorder; respiratory disorder; neurological diso
gastrointestinal disorder; urinary system disorder; drug screeni
gene therapy; chromosome mapping; forensic analysis;
antibody preparation; cytostatic; immunomodulatory; neuroprotect
antiinflammatory; gynaecological; reproductive.

Homo sapiens.

WO200200677-A1.

03-JAN-2002.

07-JUN-2001; 2001WO-US018569.

07-JUN-2000; 2000US-0209467P.

(HUMA-) HUMAN GENOME SCI INC.

Birse CE, Rosen CA;

WPI; 2002-147878/19.

N-PSDB; ABQ54751.

Isolated nucleic acid molecules encoding novel ovarian polypeptid
useful in the prevention, treatment and diagnosis of cancer (e.g
cancer), immune disorders, cardiovascular disorders and neurolog
diseases.

Claim 11; SEQ ID NO 2806; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP4
ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and al
encompasses polypeptides 90% identical and polynucleotides 95% i
to the sequences of the invention. The invention additionally re
recombinant vectors and host cells comprising human ovarian anti
polynucleotides, antibodies against human ovarian antigens, and
of ovarian antigen polynucleotides and polypeptides in diagnosin
treating, prognosing or preventing various ovary and/or breast-z
disorders. Such conditions include ovarian cancer and breast can
metastatic tumours of ovarian or breast origin, reproductive sys
disorders (e.g., infertility, disorders of pregnancy, anovulation
polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), en
disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and
shock syndrome), inflammatory conditions (e.g., mastitis, ophor
vaginits), immune disorders (e.g., congenital and acquired
immunodeficiencies, autoimmune oophoritis, systemic lupus erythe
blood-related disorders (e.g., anaemia), cardiovascular disorder
respiratory disorders, neurological disorders, gastrointestinal
and urinary system disorders. Ovarian antigen polypeptides and
polynucleotides may also be used in screening for compounds whic
modulate ovarian antigen expression or activity. The polynucleoti
further be used for gene therapy, chromosome mapping, in the
identification of individuals and in forensic analysis, and the
polypeptides may be used as food additives or to prepare antibodi
useful in disease diagnosis, drug targeting and phenotyping. The
sequence represents a human ovarian antigen of the invention. Not
specification data for this patent did not form part of the printed
specification, but was obtained in electronic format directly fr
at ftp.wipo.int/pub/published_pct_sequences

Sequence 222 AA;

Query Match 2.8%; Score 8; DB 5; Length 222;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; C

QY 60 ALGLGLAL 67
|||||

DB 190 ALGLGLAL 197
|||||

dard; protein; 286 AA.
 (first entry)
 agnostic protein #16270.
 some mapping; gene mapping; gene therapy; forensic;
 nt; medical imaging; diagnostic; genetic disorder.
 2.
 2001WO-US008631.
 2000US-00540217.
 2000US-00649167.
 INC.
 Liu C, Tang YT;
 362/73.
 166.
 polynucleotide and encoded polypeptides, useful in
 forensics, gene mapping, identification of mutations
 or genetic disorders or other traits and to assess
 ID NO 46638; 103pp; English.
 relates to isolated polynucleotide (I) and polypeptide (II)
 is useful as hybridisation probes, polymerase chain
 reaction (PCR) primers, oligomers, and for chromosome and gene ma
 inant production of (II). The polynucleotides are also used
 as expressed sequence tags for identifying expressed
 useful in gene therapy techniques to restore normal
 (II) or to treat disease states involving (II). (II) is
 useful for generating antibodies against it, detecting or quantitat
 ing tissue, as molecular weight markers and as a food
 supplement. (II) and its binding partners are useful in medical i
 maging (II). (I) and (II) are useful for treating disorders
 involving aberrant protein expression or biological activity. The
 id polynucleotide sequences have applications in
 forensics, gene mapping, identification of mutations
 or genetic disorders or other traits to assess biodiversity
 other types of data and products dependent on DNA and
 sequences. ABG0010-ABG30377 represent novel human dia
 gnoses of the invention. Note: The sequence data for this
 patent did not appear in the printed specification, but was obtained
 electronically from WIPO at
 mat directly from WIPO at
 ftp://pub.int/pub/published_pct_sequences
 A;
 2.8%; Score 8; DB 4; Length 286;
 trity 100.0%; Pred. No. 1.1e+02;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 GSS 258
 ||||
 GSS 76
 ABG01186
 ID ABG01186 standard; protein; 307 AA.
 XX
 AC ABG01186;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #1177.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR N-PSDB; AAS65373.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 31545; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypept
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene ma
 CC and in recombinant production of (II). The polynucleotides are al
 CC in diagnostics as expressed sequence tags for identifying expres
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II)
 CC useful for generating antibodies against it, detecting or quantit
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical i
 CC of sites expressing (II). (I) and (II) are useful for treating di
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodi
 CC and to produce other types of data and products dependent on DNA
 CC amino acid sequences. ABG0010-ABG30377 represent novel human dia
 CC amino acid sequences of the invention. Note: The sequence data fo
 CC patent did not appear in the printed specification, but was obtain
 CC electronic format directly from WIPO at
 CC ftp://pub.int/pub/published_pct_sequences
 XX
 SQ Sequence 307 AA;
 Query Match 2.8%; Score 8; DB 4; Length 307;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G
 QY 77 VSLGSRAS 84
 |||||
 DB 3 VSLGSRAS 10
 RESULT 54
 ADB79952
 ID ADB79952 standard; protein; 342 AA.

KQ
AA
AE
AF
AG
AH
AI
AJ
AK
AL
AM
AN
AO
AP
AQ
AR
AS
AT
AU
AV
AW
AX
AY
AZ
BA
BB
BC
BD
BE
BF
BG
BH
BI
BJ
BK
BL
BM
BN
BO
BP
BQ
BR
BS
BT
BU
BV
BW
BX
BY
BZ
CA
CB
CC
CD
CE
CF
CG
CH
CI
CJ
CK
CL
CM
CN
CO
CP
CQ
CR
CS
CT
CU
CV
CW
CX
CY
CZ
DA
DB
DC
DD
DE
DF
DG
DH
DI
DJ
DK
DL
DM
DN
DO
DP
DQ
DR
DS
DT
DU
DV
DW
DX
DY
DZ
EA
EB
EC
ED
EE
EF
EG
EH
EI
EJ
EK
EL
EM
EN
EO
EP
EQ
ER
ES
ET
EU
EV
EW
EX
EY
EZ
FA
FB
FC
FD
FE
FF
FG
FH
FI
FJ
FK
FL
FM
FN
FO
FP
FQ
FR
FS
FT
FU
FV
FW
FX
FY
FZ
GA
GB
GC
GD
GE
GF
GG
GH
GI
GJ
GK
GL
GM
GN
GO
GP
GQ
GR
GS
GT
GU
GV
GW
GX
GY
GZ
HA
HB
HC
HD
HE
HF
HG
HH
HI
HJ
HK
HL
HM
HN
HO
HP
HQ
HR
HS
HT
HU
HV
HW
HX
HY
HZ
IA
IB
IC
ID
IE
IF
IG
IH
II
IJ
IK
IL
IM
IN
IO
IP
IQ
IR
IS
IT
IU
IV
IW
IX
IY
IZ
JA
JB
JC
JD
JE
JF
JG
JH
JI
JJ
JK
JL
JM
JN
JO
JP
JQ
JR
JS
JT
JU
JV
JW
JX
JY
JZ
KA
KB
KC
KD
KE
KF
KG
KH
KI
KJ
KK
KL
KM
KN
KO
KP
KQ
KR
KS
KT
KU
KV
KW
KX
KY
KZ
LA
LB
LC
LD
LE
LF
LG
LH
LI
LJ
LK
LL
LM
LN
LO
LP
LQ
LR
LS
LT
LU
LV
LW
LX
LY
LZ
MA
MB
MC
MD
ME
MF
MG
MH
MI
MJ
MK
ML
MN
MO
MP
MQ
MR
MS
MT
MU
MV
MW
MX
MY
MZ
NA
NB
NC
ND
NE
NF
NG
NH
NI
NJ
NK
NL
NM
NO
NP
NQ
NR
NS
NT
NU
NV
NW
NX
NY
NZ
OA
OB
OC
OD
OE
OF
OG
OH
OI
OJ
OK
OL
OM
ON
OO
OP
OQ
OR
OS
OT
OU
OV
OW
OX
OY
OZ
PA
PB
PC
PD
PE
PF
PG
PH
PI
PJ
PK
PL
PM
PN
PO
PP
PQ
PR
PS
PT
PU
PV
PW
PX
PY
PZ
QA
QB
QC
QD
QE
QF
QG
QH
QI
QJ
QK
QL
QM
QN
QO
QP
QQ
QR
QS
QT
QU
QV
QW
QX
QY
QZ
RA
RB
RC
RD
RE
RF
RG
RH
RI
RJ
RK
RL
RM
RN
RO
RP
RQ
RR
RS
RT
RU
RV
RW
RX
RY
RZ
SA
SB
SC
SD
SE
SF
SG
SH
SI
SJ
SK
SL
SM
SN
SO
SP
SQ
SR
SS
ST
SU
SV
SW
SX
SY
SZ
TA
TB
TC
TD
TE
TF
TG
TH
TI
TJ
TK
TL
TM
TN
TO
TP
TQ
TR
TS
TT
TU
TV
TW
TX
TY
TZ
UA
UB
UC
UD
UE
UF
UG
UH
UI
UJ
UK
UL
UM
UN
UO
UP
UQ
UR
US
UT
UU
UV
UW
UX
UY
UZ
VA
VB
VC
VD
VE
VF
VG
VH
VI
VJ
VK
VL
VM
VN
VO
VP
VQ
VR
VS
VT
VU
VV
VW
VX
VY
VZ
WA
WB
WC
WD
WE
WF
WG
WH
WI
WJ
WK
WL
WM
WN
WO
WP
WQ
WR
WS
WT
WU
WV
WW
WX
WY
WZ
XA
XB
XC
XD
XE
XF
XG
XH
XI
XJ
XK
XL
XM
XN
XO
XP
XQ
XR
XS
XT
XU
XV
XW
XX
XY
XZ
YA
YB
YC
YD
YE
YF
YG
YH
YI
YJ
YK
YL
YM
YN
YO
YP
YQ
YR
YS
YT
YU
YV
YW
YX
YY
YZ
ZA
ZB
ZC
ZD
ZE
ZF
ZG
ZH
ZI
ZJ
ZK
ZL
ZM
ZN
ZO
ZP
ZQ
ZR
ZS
ZT
ZU
ZV
ZW
ZX
ZY
ZZ

(first entry)

BB 1 progression enhanced protein, SEQ ID 192.

ain; streptozocin-induced diabetes; rat.

gicus.

2002EP-00255249.

2001GB-00018354.

2002GB-00002910.

ER LAMBERT CO.

A, Dixon AK, Lee K, Pinnock RD;

5407/38.

9953.

ed gene sequences and encoded polypeptides that are
in the spinal cord in response to streptozocin-induced
screening compounds for the treatment of pain, or for
ain.

326-327; 334pp; English.

vention relates to nucleotide sequences which are useful in
of compounds for the treatment of pain, or for the
pain. The nucleotide sequences are up-regulated in the
in response to streptozocin-induced diabetes. The present
used to illustrate the invention.

AA;

2.8%; Score 8; DB 7; Length 342;

arity 100.0%; Pred. No. 1.3e+02;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GLAL 67

|||||

GLAL 317

dard; protein; 370 AA.

(first entry)

diagnostic protein #13382.

some mapping; gene therapy; forensic;
nt; medical imaging; diagnostic; genetic disorder.

12.

2001WO-US008631.

2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

DR N-PSDB; AAS7578.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutation
responsible for genetic disorders or other traits and to assess
biodiversity.

Claim 20; SEQ ID NO 43750; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypep
sequences. (I) is useful as hybridisation probes, polymerase cha
reaction (PCR) primers, oligomers, and for chromosome and gene m
and in recombinant production of (II). The polynucleotides are a
in diagnostics as expressed sequence tags for identifying expres
genes. (I) is useful in gene therapy techniques to restore norma
activity of (II) or to treat disease states involving (II). (II)
useful for generating antibodies against it, detecting or quanti
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical
of sites expressing (II). (I) and (II) are useful for treating d
involving aberrant protein expression or biological activity. Th
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutation
responsible for genetic disorders or other traits to assess biod
and to produce other types of data and products dependent on DNA
amino acid sequences. ABG00010-ABG30377 represent novel human di
patent did not appear in the printed specification, but was obta
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 370 AA;

Query Match

Best Local Similarity 2.8%; Score 8; DB 4; Length 370;

Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 63 LGALACL 70

|||||

Db 28 LGALACL 35

RESULT 56

ABG05012

ID ABG05012 standard; protein; 370 AA.

XX

AC ABG05012;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #5003.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

XX

PR 23-AUG-2000; 2000US-00649167.

XX

INC.

Liu C, Tang YT;

362/73.

199.

polynucleotide and encoded polypeptides, useful in diagnostics, gene mapping, identification of mutations or genetic disorders or other traits and to assess

ID NO 35371; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating it in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for patent did not appear in the printed specification, but was obtained from electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

A;

Query Match 2.8%; Score 8; DB 4; Length 370;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 0; Mismatches 0; Indels 0; Gaps 0;

ACL 70

|||

ACL 35

ard; protein; 370 AA.

first entry)

agnostic protein #18106.

ome mapping; gene mapping; gene therapy; forensic; t; medical imaging; diagnostic; genetic disorder.

001WO-US008631.

000US-00540217.

000US-00649167.

INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

DR N-PSDB; AAS82302.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 20; SEQ ID NO 48474; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide sequences (II) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating it in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for patent did not appear in the printed specification, but was obtained from electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 370 AA;

Query Match

2.8%; Score 8; DB 4; Length 370;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGLALACL 70

|||||

Db 28 LGLALACL 35

RESULT 58

ABU36537

ID ABU36537 standard; protein; 372 AA.

XX AC ABU36537;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #22064.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug de

XX OS Mycobacterium tuberculosis.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX

adio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 vick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 926/02.
 9407.
 a nucleic acids, useful for identifying proteins or screening
 as nucleic acids required for cellular proliferation to
 idate molecules for rational drug discovery programs.
) ID NO 64461; 1766pp; English.
 i relates to an isolated nucleic acid comprising any one of
 sense sequences given in the specification where expression
 c acid inhibits proliferation of a cell. Also included are:
 comprising a promoter operably linked to the nucleic acid
 peptide whose expression is inhibited by the antisense
 (2) a host cell containing the vector; (3) an isolated
 or its fragment whose expression is inhibited by the
 leic acid; (4) an antibody capable of specifically binding
 de; (5) producing the polypeptide; (6) inhibiting cellular
 i or the activity of a gene in an operon required for
 i; (7) identifying a compound that influences the activity of
 duct or that has an activity against a biological pathway
 proliferation, or that inhibits cellular proliferation; (8)
 a gene required for cellular proliferation or the biological
 which a proliferation-required gene or its gene product lies
 which the test compound that inhibits proliferation of an
 ; (9) manufacturing an antibiotic; (10) profiling a
 tivity; (11) a culture comprising strains in which the gene
 repressed or underexpressed; (12) determining the extent
 of the strains is present in a culture or collection of
 13) identifying the target of a compound that inhibits the
 i of an organism. The antisense nucleic acids are useful for
 proteins or screening for homologous nucleic acids required
 proliferation to isolate candidate molecules for rational
 y programs, or for screening homologous nucleic acids
 proliferation in cells other than *S. aureus*, *S. typhimurium*,
 or *P. aeruginosa*. The present sequence is encoded by one of
 okaryotic essential genes. Note: The sequence data for this
 t form part of the printed specification, but was obtained
 format directly from WIPO at
 pub/published_pct_sequences

AA;

2.8%; Score 8; DB 6; Length 372;
 arity 100.0%; Pred. No. 1.4e+02;
 onservative 0; Mismatches 0; Indels 0; Gaps 0;

LALA 68
 ||||
 LALA 103

dard; protein; 424 AA.

(first entry)

diagnostic protein #15604.

some mapping; gene mapping; gene therapy; forensic;
 nt; medical imaging; diagnostic; genetic disorder.

2.

XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 DR N-PSDB; AAS79800.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutation;
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 45972; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and polypep
 XX sequences. (I) is useful as hybridisation probes, polymerase cha
 CC reaction (PCR) primers, oligomers, and for chromosome and gene m
 CC and in recombinant production of (II). The polynucleotides are a
 CC in diagnostics as expressed sequence tags for identifying expres
 CC genes. (I) is useful in gene therapy techniques to restore norma
 CC activity of (II) or to treat disease states involving (II). (II)
 CC useful for generating antibodies against it, detecting or quantiti
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical
 CC of sites expressing (II). (I) and (II) are useful for treating d
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutation
 CC responsible for genetic disorders or other traits to assess biod
 CC and to produce other types of data and products dependent on DNA
 CC amino acid sequences. ABG00010-ABG30377 represent novel human di
 CC amino acid sequences of the invention. Note: The sequence data f
 CC patent did not appear in the printed specification, but was obta
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 424 AA;
 SQ

Query Match 2.8%; Score 8; DB 4; Length 424;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; G

OY 60 ALGGLAL 67
 |||||
 Db 392 ALGGLAL 399

RESULT 60
 AAM23752
 ID AAM23752 standard; protein; 430 AA.
 XX
 AC AAM23752;
 XX

12-OCT-2001 (first entry)

Human EST encoded protein SEQ ID NO: 1277.

Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse
 tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 diagnostics; forensic test; gene mapping; genetic disorder; biodi
 gene therapy; nutrition.

Homo sapiens.

WO200154477-A2.

02-AUG-2001.

2001WO-US002687.

2000US-00491404.

2000US-00617746.

2000US-00631451.

2000US-00663870.

INC.

C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
ac RA, Zhang J, Werhman T;

164/51.

411.

peptide for treatment of diseases, diagnostics, raising
a research use.

2 920; 1275pp; English.

vention provides the protein and coding sequences of novel
a variety of organisms, including human, dog, cat, horse,
ster, monkey, macaque, yeast, bacteria, fruit fly, sea
nato. These were derived from expressed sequence tags (ESTs)
ism of interest. They can be used in diagnostics,
e mapping, identification of mutations, to assess
and for nutritional purposes. The present sequence is a
invention

AA;

2.8%; Score 8; DB 4; Length 430;

100.0%; Pred. No. 1.6e+02;

nservative 0; Mismatches 0; Indels 0; Gaps 0;

AL 61

|||

AL 9

hard; protein; 431 AA.

first entry)

sequence Seq ID508 related to grain filling.

ology; carbohydrate synthesis; carbohydrate metabolism;
egradation; carbohydrate; plant grain; grain filling; corn;
; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
rotein; oil; starch; fibre; moisture content; cereal grain;
t.

2.

002WO-IB002450.

001US-0300112P.

001US-032527P.

001US-0342327P.

TA PARTICIPATIONS AG.

W, Briggs S, Cooper B, Goff SA, Moughamer T;
Katagiri F, Kreps J, Provart N, Ricke D;

XX

WPI; 2003-229341/22.

N-PSDB; ADC08202.

XX

New plant genes encoding polypeptides having an activity involve
associated with the synthesis, metabolism or degradation of carb
in the plant grain useful in generating plants having improved
nutritional properties.

PT

Claim 34; SEQ ID NO 508; 130pp; English.

XX

This invention, in the area of plant biotechnology, relates to ne
polynucleotides comprising a nucleotide sequence encoding a prote
is involved in or associated with the synthesis, metabolism or
degradation of carbohydrates in the plant grain and the expres
which is up-regulated during grain filling. The plant is selected
corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
sugarbeet, wheat, and rice. The invention may be useful for the
improvement of protein, oil, starch, fibre and moisture content c
cereal grains. In addition, carbohydrate levels may be modified t
desirable level using the present invention. The present sequence
amino acid sequence of a rice protein of the invention. Note: The
sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly fro
at ftp.wipo.int/pub/publishedpct_sequences.

XX SQ Sequence 431 AA;

Query Match 2.8%; Score 8; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 42 QRRGRGG 49

Db 372 QRRGRGG 379

RESULT 62

ADC64562

ID ADC64562 standard; protein; 431 AA.

XX

AC ADC64562;

XX

01-JAN-2004 (first entry)

XX

Synechococcus sp. Synwh0268 protein.

DE

Plant growth; commercial yield; plant breeding; fruit yield;

XX

flowering rate; Synwh0268.

KW

Synechococcus sp.; WH 8102.

XX

US2003192076-A1.

XX

09-OCT-2003.

XX

10-APR-2003; 2003US-00410432.

XX

26-MAR-2002; 2002WO-IL000250.

XX

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX

Kaplan A, Lieman-Hurwitz J, Schatz D, Mittler R, Rachmilevitch

XX

WPI; 2003-831832/77.

XX

Obtaining plants having enhanced growth and/or fruit yield and/or

PT

flowering rate, specifically C3 plants grown under limiting condit

XX

useful in plant molecular biology and commercial plant breeding.

XX

Claim 2; Fig 11; 47pp; English.

PS

The present invention relates to a method of obtaining plants with

XX

with and/or commercial yield under growth limiting conditions. comprises obtaining a population of plants transformed to a polypeptide having at least 60% sequence identity to any of 8 sequences, growing the plants and selecting plants having the polypeptide. The methods and compositions of the present invention are useful in commercial plant breeding, particularly for plants having enhanced growth and/or fruit yield and/or ornamental value. The present sequence represents *Synechococcus* sp. protein.

AA;

2.8%; Score 8; DB 7; Length 431;
 Identity 100.0%; Pred. No. 1.6e+02;
 Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LG LLL 74
 |||||
 LG LLL 371

standard; protein; 454 AA.

(first entry)

elanogaster polypeptide SEQ ID NO 516.

developmental biology; cell signalling; insecticide;
 al.

elanogaster.

A2.

2001WO-US009231.

2000US-0191637P.

2000US-00614150.

DRP NY.

Adams M, Li PWD, Myers EW;

5860/75.

2011.

nucleic acid detection reagent for detecting 1000 or more
 Drosophila and for elucidating cell signalling and cell-cell

SEQ ID NO 516; 21pp + Sequence Listing; English.

relates to an isolated nucleic acid detection reagent
 detecting 1000 or more genes from *Drosophila*. The invention is
 experimental biology and in elucidating cell signalling and
 interactions in higher eukaryotes for the development of
 therapeutics and pharmaceutical drugs. The invention
 genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 H01840-ABL16175) and the encoded proteins (ABB57737-
 the sequence data for this patent did not form part of the
 publication, but was obtained in electronic format directly
 ftp.wipo.int/pub/published_pct_sequences

AA;

2.8%; Score 8; DB 4; Length 454;
 Identity 100.0%; Pred. No. 1.7e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0;
 QY 56 LVPLALGL 63
 |||||
 Db 170 LVPLALGL 177
 |||||

RESULT 64

ADA54710
 ID ADA54710 standard; protein; 472 AA.

XX

AC ADA54710;

XX

DT 20-NOV-2003 (first entry)

XX

DE Human protein, SEQ ID 2278.

XX

KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Noc
 Gene Therapy; human; secretory protein; membrane proteins; cancer
 inflammatory disease; osteoporosis; neurological disease.

XX

OS Homo sapiens.

XX

PN EP1293569-A2.

XX

PD 19-MAR-2003.

XX

PF 21-MAR-2002; 2002EP-00006586.

XX

PR 14-SEP-2001; 2001JP-00328381.

XX

PR 24-JAN-2002; 2002US-0350435P.

XX

(HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii
 Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tam
 Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
 WPI: 2003-395539/38.
 DR N-PSDB; ADA53071.

XX

PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 and/or membrane proteins, useful for developing medicines for di
 PT which the gene is involved, or as target molecules for gene ther.
 XX

PS

Claim 14; SEQ ID NO 2278; 205pp; English.

XX

CC The present invention relates to novel human secretory or membra
 proteins (ADA54072-ADA55710) and their coding sequences (ADA5243
 CC ADA54071). The coding sequences are useful in the gene therapy o
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX

SQ

Sequence 472 AA;

Query Match

Best Local Similarity 2.8%; Score 8; DB 6; Length 472;

Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY

60 ALGLGLAL 67

|||||

Db 440 ALGLGLAL 447

|||||

RESULT 65

ABG20260

ID ABG20260 standard; protein; 586 AA.

XX

AC ABG20260;

XX

DT 18-FEB-2002 (first entry)

XX

diagnostic protein #20251.

some mapping; gene mapping; gene therapy; forensic;
nt; medical imaging; diagnostic; genetic disorder.

2.

2001WO-US008631.

2000US-00540217.

2000US-00649167.

INC.

Liu C, Tang YT;

362/73.

447.

polynucleotide and encoded polypeptides, useful in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits and to assess

ID NO 50619; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
is useful as hybridisation probes, polymerase chain
primers, oligomers, and for chromosome and gene mapping,
inant production of (II). The polynucleotides are also used
s as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal
II) or to treat disease states involving (II). (II) is
nerating antibodies against it, detecting or quantitating a
tissue, as molecular weight markers and as a food
II) and its binding partners are useful in medical imaging
essing (II). (I) and (II) are useful for treating disorders
rant protein expression or biological activity. The
ad polynucleotide sequences have applications in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits to assess biodiversity
e other types of data and products dependent on DNA and
quences. ABG0010-ABG30377 represent novel human diagnostic
ences of the invention. Note: The sequence data for this
appear in the printed specification, but was obtained in
mat directly from WIPO at
pub/published_pct_sequences

LA;

2.8%; Score 8; DB 4; Length 586;
arity 100.0%; Pred. No. 2.2e+02;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

ALA 68

ALA 21

iard; protein; 586 AA.

(first entry)

P36269, SEQ ID NO 11960.

Human; pain; neuronal tissue; gene therapy;
spinal segmental nerve injury; chronic constriction injury; CCI;
spared nerve injury; SNI; Chung.

Homo sapiens.

WO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002WO-US025765.

14-AUG-2001; 2001US-0312147P.

01-NOV-2001; 2001US-0346382P.

26-NOV-2001; 2001US-0333347P.

(GEHO) GEN HOSPITAL CORP.

(FARB) BAYER AG.

Wolff C, D'urso D, Befort K, Costigan M;

WPI; 2003-268312/26.

GENBANK; P36269.

New composition comprising two or more isolated polypeptides, use
preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isol
or human polynucleotides or a polynucleotide which represents a f
derivative or allelic variation of the nucleic acid sequence. Als
claimed are a vector comprising the novel polynucleotide, a host
comprising the vector, a method for identifying a nucleotide seq
which is differentially regulated in an animal subjected to pain
kit to perform the method, an array, a method for identifying an
that increases or decreases the expression of the polynucleotide
that is differentially expressed in neuronal tissue of a first ar
subjected to pain, a method for identifying a compound which regu
the expression of a polynucleotide sequence which is differentia
expressed in an animal subjected to pain, a method for identifyin
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composi
method for identifying a compound or small molecule that regulate
activity in an animal of one or more of the polypeptides given in
specification, a method for identifying a compound useful in trea
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compo
modulates its activity is useful for preparing a medicament for t
pain (e.g. spinal segmental nerve injury (SNI)), chronic constri
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. ge
therapy). The sequence presented is a human protein (shown in Tab
the specification) which is differentially expressed during pain.
The sequence data for this patent did not form part of the printe
specification, but was obtained in electronic form directly from
ftp.wipo.int/pub/published_pct_sequences.

Sequence 586 AA;

Query Match 2.8%; Score 8; DB 7; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; G

QY 61 LGUGLALA 68

Db 14 LGUGLALA 21

RESULT 67

ADE62980

ID ADE62980 standard; protein; 586 AA.

XX AC ADE62980;

(first entry)

n P36269, SEQ ID NO 8914.

neural tissue; gene therapy;
 neural nerve injury; chronic constriction injury; CCI;
 injury; SNI; Chung.

-A2.

2002WO-US025765.

2001US-0312147P.

2001US-0346382P.

2001US-0333347P.

HOSPITAL CORP.

R AG.

urso D, Befort K, Costigan M;

8312/26.

269.

ion comprising two or more isolated polypeptides, useful for
 medicament for treating pain in an animal.

e; 1017pp; English.

n discloses a composition comprising two or more isolated rat
 nucleotides or a polynucleotide which represents a fragment,
 r allelic variation of the nucleic acid sequence. Also
 a vector comprising the novel polynucleotide, a host cell
 e vector, a method for identifying a nucleotide sequence
 ferentially regulated in an animal subjected to pain and a
 m the method, an array, a method for identifying an agent
 es or decreases the expression of the polynucleotide sequence
 nentially expressed in neuronal tissue of a first animal
 pain, a method for identifying a compound which regulates
 on of a polynucleotide sequence which is differentially
 an animal subjected to pain, a method for identifying a
 regulates the activity of one or more of the
 ies, a method for producing a pharmaceutical composition, a
 identifying a compound or small molecule that regulates the
 an animal of one or more of the polypeptides given in the
 a, a method for identifying a compound useful in treating
 armaceutical composition comprising the one or more
 or their antibodies. The polynucleotide or the compound that
 activity is useful for preparing a medicament for treating
 inal segmental nerve injury (Chung), chronic constriction
 and spared nerve injury (SNI)) in an animal (e.g. gene
 sequence presented is a human protein (shown in Table 2 of
 ition) which is differentially expressed during pain. Note:
 data for this patent did not form part of the printed
 l, but was obtained in electronic form directly from WIPO at
 /pub/published_pct_sequences.

AA;

2.8%; Score 8; DB 7; Length 586;
 arity 100.0%; Pred. No. 2.2e+02;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

HLALA 68

||||

HLALA 21

RESULT 68
 AAU32148
 ID AAU32148 standard; protein; 592 AA.
 XX
 AC AAU32148;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #2639.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regenerati
 KW immune suppression; immune stimulation; anti-inflammatory; leuk
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US008656.
 XX
 PR 18-APR-2000; 2000US-00552929.
 PR 28-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in
 PT vaccination, testing and therapy.
 XX
 PS Claim 20; Page 562-563; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease assoc
 CC with altered levels of polypeptide. The polypeptides are also us
 CC identifying agents (agonists and antagonists) that bind to them.
 CC expressing the proteins are useful for identifying a therapeutic
 CC for use in treatment of a pathology related to aberrant expressi
 CC physiological interactions of the polypeptide. Vectors compris
 CC nucleic acids encoding the polypeptides and cells genetically en
 CC to express them are also useful for producing the proteins. The
 CC are useful in genetic vaccination, testing and therapy, and can
 CC as nutritional supplements. They may be used to increase stem ce
 CC proliferation; to regulate haematopoiesis; and in bone, cartilag
 CC and/or nerve tissue growth or regeneration; immune suppression a
 CC stimulation; as anti-inflammatory agents; and in treatment of le
 CC AAU29510-AAU3304 represent the amino acid sequences of novel hu
 CC secreted proteins of the invention
 XX
 SQ Sequence 592 AA;
 Query Match 2.8%; Score 8; DB 4; Length 592;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;
 QY 61 LGLGLALA 68
 |||||
 DB 44 LGLGLALA 51
 RESULT 69
 ABG03722
 ID ABG03722 standard; protein; 603 AA.
 XX
 AC ABG03722;
 XX
 DT 13-FEB-2002 (first entry)
 XX

agnostic protein #3713.
 some mapping; gene mapping; gene therapy; forensic;
 it; medical imaging; diagnostic; genetic disorder.
 1.
 ?001WO-0008631.
 ?000US-00540217.
 ?000US-00649167.
 INC.
 Liu C, Tang YT;
 162/73.
 109.
 polynucleotide and encoded polypeptides, useful in
 forensics, gene mapping, identification of mutations
 or genetic disorders or other traits and to assess
 ID NO 34081; 103pp; English.
 relates to isolated polynucleotide (I) and polypeptide (II)
 is useful as hybridisation probes, polymerase chain
 primers, oligomers, and for chromosome and gene mapping,
 nant production of (II). The polynucleotides are also used
 as expressed sequence tags for identifying expressed
 useful in gene therapy techniques to restore normal
 I) or to treat disease states involving (II). (II) is
 erating antibodies against it, detecting or quantitating a
 tissue, as molecular weight markers and as a food
 I) and its binding partners are useful in medical imaging
 ssing (II). (I) and (II) are useful for treating disorders
 rant protein expression or biological activity. The
 d polynucleotide sequences have applications in
 onresics, gene mapping, identification of mutations
 r genetic disorders or other traits to assess biodiversity
 other types of data and products dependent on DNA and
 vences. ABG00010-ABG30377 represent novel human diagnostic
 vences of the invention. Note: The sequence data for this
 appear in the printed specification, but was obtained in
 mat directly from WIPO at
 ub/published_pct_sequences
 A;
 2.8%; Score 8; DB 4; Length 603;
 rity 100.0%; Pred. No. 2.3e+02;
 nservative 0; Mismatches 0; Indels 0; Gaps 0;
 ALA 68
 |||
 ALA 447
 ard; protein; 617 AA.
 first entry)
 modification and maintenance molecule (PMMW)-44.

KW protein modification and maintenance molecule; PMMW;
 KW protein modification; protein maintenance; protein function;
 KW protein conformation; protein stabilisation; protein degradation;
 KW phosphatase; protease; protease inhibitor; isomerase; transferase
 KW molecular chaperone; anti-HIV; anti-allergic; anti-inflammatory;
 KW antianaemic; antiparkinsonian; neurotropic; anticonvulsant;
 KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyro
 KW cytostatic; hepatotropic; dermatological; antidiabetic; nephro
 KW antitumor; thyromimetic; neuroprotective; osteopathic; antiarthritis
 KW antiparasitic; antihelminthic; antipeptidic; uropathic; ophthalmic
 KW antineumatic; haemostatic; antibacterial; virucide; protozoacide
 KW fungicide; gene therapy; cell proliferative disorder; arterioscle
 KW hepatitis; polycythaemia vera; psoriasis; primary thrombocytopen
 KW cancer; developmental disorder; anaemia; mental retardation;
 KW neurological disorder; Alzheimer's disease; Parkinson's disease;
 KW epilepsy; autoimmune disorder; inflammatory disorder; AIDS; aller
 KW asthma; autoimmune thyroiditis; Crohn's disease; diabetes mellitu
 KW glomerulonephritis; Goodpasture's syndrome; multiple sclerosis;
 KW arthritis; osteoporosis; pancreatitis; Sjogren's syndrome;
 KW microbial infection; human.
 XX Homo sapiens.
 OS
 XX
 XX
 PN WO2003063688-A2.
 XX
 XX
 PD 07-AUG-2003.
 XX
 XX
 PF 23-JAN-2003; 2003WO-US002500.
 XX
 XX
 PR 25-JAN-2002; 2002US-0351928P.
 PR 25-FEB-2002; 2002US-0359903P.
 PR 21-MAR-2002; 2002US-0366837P.
 XX
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Hafalia AJA, Li JX, Gorvad AE, Chawla NK, Sprague WW, Lee SY,
 PI Chang H, Elliott VS, Ramkumar J, Khare R, Emerling BM, Kable
 PI Tang YT, Yue H, Gietzen KJ, Lee S, Swarnakar A, Baughn WR;
 PI Wilson AD, Jin P, Chien D, Hawkins PR, Jiang X, Jackson AA;
 PI Bhadia U, Burrill JD, Blake JJ, Ho A, Zheng W, Ison CH, Marc
 PI Tran UK, Lal PG, Warren BA, Xu Y, Honchell CD, Becha SD;
 PI Lehr-Mason PM;
 XX
 DR WPI; 2003-636761/60.
 DR N-PSDB; ADE79064.
 XX
 XX
 PT New human protein modification and maintenance molecules and
 PT polynucleotides, useful for diagnosing, treating or preventing aut
 PT or inflammatory disorders (e.g. AIDS, allergy or anemia), multiple
 PT sclerosis or cancer.
 XX
 PS Claim 1; SEQ ID NO 44; 405pp; English.
 XX
 CC This invention relates to novel isolated human proteins, which are
 CC protein modification and maintenance molecules (PMMW). The cellula
 CC processes regulating modification and maintenance of protein molec
 CC coordinate their function, conformation, stabilisation and degrada
 CC Each of these processes is mediated by key enzymes or proteins suc
 CC kinases, phosphatases, proteases, protease inhibitors, isomerases,
 CC transferases and molecular chaperones. Compounds which modulate th
 CC proteins of the invention may have anti-HIV, anti-allergic,
 CC antiinflammatory, antianaemic, antiparkinsonian, neurotropic,
 CC anticonvulsant, antiarteriosclerotic, antiasthmatic, immunosuppres
 CC antithyroid, cytostatic, hepatotropic, dermatological, antidiabeti
 CC nephrotropic, antitumor, thyromimetic, neuroprotective, osteopathi
 CC antiarthritis, antiparasitic, antihelminthic, antipeptidic, uropat
 CC ophthalmological, antirheumatic, haemostatic, antibacterial, viruc
 CC protozoacide or fungicide activities. The DNA sequence which encod
 CC proteins of the invention may be useful for gene therapy. The huma
 CC protein modification and maintenance molecules (PMMWs), the DNA se
 CC which encode them and their modulating compounds are useful for
 CC diagnosing, treating or preventing disorders associated with aberr
 CC expression of PMMW, particularly cell proliferative disorders (for

riosclerosis, hepatitis, polycythemia vera, psoriasis, monocytopenia or cancer), developmental disorders (for mia or mental retardation), neurological disorders (for eimer's disease, Parkinson's disease or epilepsy), inflammatory disorders (for example AIDS, allergies, asthma, hyroiditis, Crohn's disease, diabetes mellitus, hritis, Goodpasture's syndrome, multiple sclerosis, steoporosis, pancreatitis, Sjogren's syndrome) or microbial The present sequence is the amino acid sequence of a human invention.

AA;

2.8%; Score 8; DB 7; Length 617;
larity 100.0%; Pred. No. 2.3e+02;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3LALA 68

|||||

3LALA 21

ardard; protein; 633 AA.

(first entry)

diagnostic protein #20252.

osome mapping; gene mapping; gene therapy; forensic;
ant; medical imaging; diagnostic; genetic disorder.

A2.

2001WO-US008631.

2000US-00540217.

2000US-00649167.

) INC.

Liu C, Tang YT;

362/73.

448.

polynucleotide and encoded polypeptides, useful in
forensics, gene mapping, identification of mutations
or genetic disorders or other traits and to assess

ID NO 50620; 103pp; English.

relates to isolated polynucleotide (I) and polypeptide (II)
is useful as hybridisation probes, polymerase chain
primers, oligomers, and for chromosome and gene mapping,
inant production of (II). The polynucleotides are also used
s as expressed sequence tags for identifying expressed
useful in gene therapy techniques to restore normal
II) or to treat disease states involving (II). (II) is
nerating antibodies against it, detecting or quantitating a
n tissue, as molecular weight markers and as a food
II) and its binding partners are useful in medical imaging
essing (II). (I) and (II) are useful for treating disorders
rrent protein expression or biological activity. The
nd polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutation
CC responsible for genetic disorders or other traits to assess bio
CC and to produce other types of data and products dependent on DN
CC amino acid sequences. ABC0010-ABG30377 represent novel human di
CC amino acid sequences. ABC0010-ABG30377 represent novel human di
CC amino acid sequences. ABC0010-ABG30377 represent novel human di
CC patent did not appear in the invention. Note: The sequence data f
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 633 AA;

Query Match

Best Local Similarity 2.8%; Score 8; DB 4; Length 633;

Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 61 LGLGLALA 68

|||||

Db 44 LGLGLALA 51

RESULT 72

ABG28291

ID ABG28291 standard; protein; 799 AA.

XX AC ABG28291;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #28282.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS92478.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 58650; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypept
CC sequences. (I) is useful as hybridisation probes, polymerase chai
CC reaction (PCR) primers, oligomers, and for chromosome and gene ma
CC and in recombinant production of (II). The polynucleotides are al
CC in diagnostics as expressed sequence tags for identifying expres
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II)
CC useful for generating antibodies against it, detecting or quantit
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical i
CC of sites expressing (II). (I) and (II) are useful for treating di
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodi

Other types of data and products dependent on DNA and sequences. ABG00010-ABG30377 represent novel human diagnostic sequences of the invention. Note: The sequence data for this report appear in the printed specification, but was obtained in format directly from WIPO at pub/published_pct_sequences

AA;

2.8%; Score 8; DB 4; Length 799;
arity 100.0%; Pred. No. 3e+02;
nservative 0; Mismatches 0; Indels 0; Gaps 0;

ELAL 67

|||||

ELAL 294

ard; protein; 842 AA.

(first entry)

(useful for identifying genetic disorders) #630.

vel protein; tissue marker; molecular weight marker;
ker; genetic disorder.

2.

:002WO-US0393555.

:001US-0339739P.

:001US-0339453P.

:002US-0365091P.

:002US-0365384P.

:002US-0372381P.

:002US-0372615P.

:002US-00128558.

:002US-0376045P.

INC.

di V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
Chen R, Xu C, Boyle BJ;

35/53.

64.

tides, useful for expressing recombinant proteins for
acterization or therapeutic use, or as markers for tissues
corresponding protein is preferentially expressed.

ID NO 1541; 1177pp; English.

comprises the amino acid and coding sequences of novel
DNA and protein sequences of the invention are useful as:
ssues in which the corresponding protein is preferentially
molecular weight markers on gels; as chromosome markers or
ify chromosomes or to map related gene positions; and to
ndogenous DNA sequences in patients to identify potential
ers. The present amino acid sequence represents a protein
on.

A;

Query Match 2.8%; Score 8; DB 7; Length 842;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; C

Qy 54 ALLVPLAL 61

|||||

Db 441 ALLVPLAL 448

RESULT 74

ABU16705

ID ABU16705 standard; protein; 1032 AA.

XX

AC

ABU16705;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #2232.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug d

XX

OS Acinetobacter baumannii.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR

PR 06-SEP-2001; 2001US-00948993.

PR

PR 25-OCT-2001; 2001US-0342923P.

PR

PR 08-FEB-2002; 2002US-00072851.

PR

PR 06-MAR-2002; 2002US-0362699P.

XX

(ELIT-) ELITRA PHARM INC.

XX

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

WPI; 2003-029926/02.

DR

N-PSDB; ACA20575.

XX

New antisense nucleic acids, useful for identifying proteins or s
for homologous nucleic acids required for cellular proliferation
isolate candidate molecules for rational drug discovery programs.

PT

PT

XX

Claim 25; SEQ ID NO 44629; 1766pp; English.

XX

The invention relates to an isolated nucleic acid comprising any
the 6213 antisense sequences given in the specification where exp
of the nucleic acid inhibits proliferation of a cell. Also include
CC (1) a vector comprising a promoter operably linked to the nucleic
CC encoding a polypeptide whose expression is inhibited by the antis
CC nucleic acid; (2) a host cell containing the vector; (3) an isolat
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically bi
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cell
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the acti
CC the gene product or that has an activity against a biological path
CC required for proliferation, or that inhibits cellular proliferati
CC identifying a gene required for cellular proliferation or the biol
CC pathway in which a proliferation-required gene or its gene product
CC or a gene on which the test compound that inhibits proliferation c
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which th
CC product is overexpressed or underexpressed; (12) determining the e
CC to which each of the strains is present in a culture or collection
CC strains; or (13) identifying the target of a compound that inhibit
CC proliferation of an organism. The antisense nucleic acids are usef
CC identifying proteins or screening for homologous nucleic acids req

proliferation to isolate candidate molecules for rational
 ry programs, or for screening homologous nucleic acids
 proliferation in cells other than *S. aureus*, *S. typhimurium*,
 e or *P. aeruginosa*. The present sequence is encoded by one of
 rokaryotic essential genes. Note: The sequence data for this
 ot form part of the printed specification, but was obtained
 c format directly from WIPO at
 /pub/published_pct_sequences

2 AA;

2.8%; Score 8; DB 6; Length 1032;
 larity 100.0%; Pred. No. 3.8e+02;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALGLG 64

|||||

ALGLG 981

ndard; protein; 1033 AA.

(first entry)

r baumannii protein #1012.

r baumannii; bacterial disease; antibacterial; vaccine;
 crol agent.

r baumannii.

99US-00328352.

98US-0088701P.

4E THERAPEUTICS CORP.

ish D;

5092/54.

1725.

acter baumannii proteins and nucleic acids, useful as reagents
 ing a bacterial disease, as components of antibacterial
 targets for antibacterial drugs, or as biocontrol agents for

ID NO 5138; 328pp; English.

i relates to isolated Acinetobacter baumannii nucleic acids.
 uni nucleic acids and polypeptides are useful as reagents
 ing a bacterial disease, as components of antibacterial
 targets for antibacterial drugs, to detect the presence of
 and other Acinetobacter species in a sample, in screening
 the ability to interfere with the A. baumannii life cycle
 : A. baumannii infection, and as biocontrol agents for
 present sequence represents the amino acid sequence of an A.
 otein.

AA;

2.8%; Score 8; DB 6; Length 1033;

arity 100.0%; Pred. No. 3.8e+02;

Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 VPLALGLG 64
 |||||

Db 975 VPLALGLG 982

Search completed: April 7, 2004, 17:57:27
 Job time : 65 secs